(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 370 458 A2**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 89121513.9

(51) Int. Cl.5: C07K 13/00, C12N 15/62, G01N 33/569

(22) Date of filing: 21.11.89

(30) Priority: 23.11.88 US 275309

(43) Date of publication of application: 30.05.90 Bulletin 90/22

(84) Designated Contracting States: AT BE CH DE ES FR GB GR IT LI NL SE

(71) Applicant: ABBOTT LABORATORIES
One Abbott Park Road
Abbott Park, IL 60064-3500(US)

(72) Inventor: Devare, Sushil G.
2492 Farnsworth Lane
Northbrook Illinois 60062(US)
Inventor: Suresh, Desai M.
1408 Amy Lane
Libertyville Illinois 60048(US)
Inventor: Casey, James M.
4567 McClure
Gurnee Illinois 60031(US)

(74) Representative: Modiano, Guido et al
MODIANO, JOSIF, PISANTY & STAUB
Modiano & Associati Via Meravigli, 16
I-20123 Milano(IT)

(54) Synthetic DNA derived recombinant HIV antigens.

(57) The present invention provides a method of synthesizing genes encoding unique HIV-1 and HIV-2 envelope proteins and their fragments, thereby allowing overexpression of these proteins in *E. coli*. The HIV envelope proteins and their fragments have been expressed at high levels as individual proteins or in fusion with other proteins. The HIV envelope proteins thus expressed in *E. coli* can be effectively used for the detection of exposure to HIV as well as the discrimination of HIV-1 and HIV-2.

EP 0 370 458 A2

Clustered order of selected sequences:

2. CDC42FRAG.PEP (1-107)
3. BH102FRAG.PEP (1-107)
4. SF2FRAG.PEP (1-107)
1. MALFRAG.PEP (1-107)
5. SYNFRAG.PEP (1-107)

```
2     1 kAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGfWGCSGKLICTTAVPWNASWSNKtLdQIWNNMT
3     1 EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWNNMT
4     1 EAQQHLLQLTVWGIKQLQARVLAVERYLrDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEDIWdNMT
1     1 EAQQHLLQLTVWGIKQLQARVLAVERYLqDQrLLGmWGCSGKhICTTfVPWNsSWSNrSLdDIWnNMT
5     1 KAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEDIWNNMT

2    69 WMEWDREIdNYThLIytLIEESQNQQEKNqQELLqLDKW
3    69 WMEWDREInNYTsLIhsLIEESQNQQEKNEQELLELDKW
4    69 WMQWEREIdNYTntIYtLIEESQNQQEKNEQELLELDKW
1    69 WMQWEkEIsNYTgiIYnLiEESQiQQEKNEkELLELDKW
5    69 WMQWEREINNYTNLIYSLLEESQNQQEKNEQELLQLDKW
```

FIGURE 1

SYNTHETIC DNA DERIVED RECOMBINANT HIV ANTIGENS

BACKGROUND OF THE INVENTION

The present invention relates to recombinant HIV (Human Immunodeficiency Virus) antigens. Recombinant antigens derived from the molecular cloning and expression in a heterologous expression system of the synthetic DNA sequences of the various HIV antigens can be used as reagents for the detection of antibodies and antigen in body fluids from individuals exposed to various HIV isolates.

The nucleotide sequence of the proviral genome has been determined for several HIV isolates, including HIV-1 strains HTLV-III (Ratner et al., *Nature* (1985) 313:277); ARV-2 (Sanchez-Pescador et al., *Science* (1985) 227:484); LAV (Wain-Hobson et al., *Cell* (1985) 40:9); and CDC-451 (Desai et al., *Proc. Natl. Acad. Sci. USA* (1986) 83:8380). The nucleotide sequence of the HIV-2 ROD isolate was reported by Guyader et al. (*Nature* (1987) 326:662).

HIV antigens have been obtained from the virus grown in tissue culture, or from a molecularly cloned genomic fragment expressed in heterologous hosts such as Escherichia coli. The tissue culture derived virus involves the cumbersome and often difficult process of growing virus infected cells in stringent sterile conditions. Further, the virus derived from tissue culture is infectious, and, therefore is hazardous to the health of individuals involved in propagation and purification. The expression of molecularly cloned HIV genomic fragments overcomes the biohazard problem. Generally, an HIV genomic fragment from a single HIV isolate with mammalian codons is expressed in a heterologous system, such as, bacteria or yeast, and is limited to the use of available restriction sites present in the viral genome for cloning and expression.

It has been difficut to obtain expression in heterologous systems of some of the HIV proteins, such as the HIV-1 envelope antigen gp41. Several researchers have tried deleting the hydrophobic regions of the HIV-1 gp41 to increase expression levels. UK Patent Application GB 2188639 discloses an HTLV-III gag/env gene protein wherein the env fragment of the DNA sequence deleted codons corresponding to the first hydrophobic region of the gp41 protein. U.S. Patent No. 4,753,873 discloses a peptide fragment that is encoded by a nucleotide sequence wherein the nucleotides coding for a first and second hydrophobic region of HTLV-III gp41 are deleted.

Poor expression can be the result of many factors, including the specific nucleic acid sequence of the gene to be expressed, the mammalian codons of the gene sequence to be expressed may not be efficiently transcribed and translated in a particular heterologous system, and the secondary structure of the transcribed messenger RNA. The use of synthetic DNA fragments can increase expression in heterologous systems.

SUMMARY OF THE INVENTION

Recombinant antigens which are derived from the molecular cloning and expression of synthetic DNA sequences in heterologous hosts are provided. Synthetic DNA sequences coding for the recombinant antigens of the invention are further provided. The synthetic DNA sequences selected for expression of various HIV antigens are based on the amino acid sequence of either a single isolate or several isolates, optimized for expression in Escherichia coli by specific codon selection. The synthetic DNA sequence gives higher expression of the particular antigen encoded. These antigens can be substituted for viral antigens derived from tissue culture for use as diagnostic and therapeutic reagents.

The present invention can be utilized to synthesize full length HIV transmembrane envelope gene using bacterial codons. Another aspect of the invention involves the linkage of sequences which are poorly expressed as individual proteins, to sequences which are expressed with high efficiency. The combination of the sequence of the entire coding region of a gene of one virus with coding sequences of another gene from a different virus to produce a fusion protein can be achieved. The fusion proteins thus expressed have a unique advantage of antigenic epitopes of two viral antigens.

The present invention includes full length synthetic genes (FSG) for HIV-1 and HIV-2 transmembrane glycoprotein (TMP).

DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates the alignment of the TMP fragment encoding amino acid residue nos. 552-668 of HIV-1 with the sequences of the four different isolates used to derive the amino acid sequence of BS2-10.

Fig. 2 illustrates the assembly of 16 oligonucleotides to form the synthetic TMP fragment of Fig. 1, and its cloning into pUC18, designated BS2-10.

Fig. 3 illustrates the DNA and amino acid sequence of FSG, indicating the restriction sites and subfragments used for assembly.

Fig. 4 is a comparison of the amino acid sequence used to develop the synthetic HIV-1 envelope gene with known amino acid sequences of 13 independent isolates, indicating all linker-derived sequences (+) and amino acid substitutions (*).

Fig. 5 is a schematic diagram of the assembly and cloning of the major subfragments to form FSG in pUC18.

Fig. 6 is a schematic diagram of the cloning of FSG into lambda pL expression vectors to generate pSD301 and pSD302.

Figs. 7 illustrates the amino acid sequences of pSD301 and pSD302, indicating all linker-derived sequences (+) and amino acid substitutions (*).

Fig. 8 illustrates results of expression analysis of pSD301 and pSD302. A) Coomassie stained gel; B) Immunoblot using AIDS patients' sera.

Fig. 9 illustrates the DNA and amino acid sequence of the full length synthetic HIV-2 TMP, indicating restriction enzymes used to assemble the gene including linker sequences at both ends to facilitate cloning.

Fig. 10 illustrates the three major subfragments used to construct the synthetic HIV-2 TMP gene.

Fig. 11 is a schematic diagram of the assembly of the major subfragments to form the full length synthetic HIV-2 TMP and its cloning into pUC8 to generate pJC28.

Fig. 12 is a schematic diagram of the cloning of synthetic HIV-2 TMP fragment A into pUC19 to generate pJC22 and into pTB210 to generate pJC100.

Fig. 13 is a schematic diagram of the cloning of synthetic HIV-2 TMP into lambda pL expression vectors to generate pSD306 and pSD307.

Fig. 14 indicates the specific amino acid sequences of pL constructs pSD306 and pSD307, indicating all linker sequences, HIV-1 gag sequences, and HIV-2 TMP sequences.

Fig. 15 illustrates results of expression analysis of pSD306 in *E. coli* CAG456 cells. A) Coomassie stained gel; B) Immunoblot using HIV-2 positive human sera.

Fig. 16 illustrates results of expression analysis of pSD307 in *E. coli* pRK248.clts/RR1 cells. A) Coomassie stained gel; B) Immunoblot using HIV-2 positive human sera.

DETAILED DESCRIPTION OF THE INVENTION

Synthetic DNA fragments of the HIV genome can be synthesized based on their corresponding amino acid sequences. By comparing the particular region of interest between different isolates, a sequence can be selected which is different from any sequence that exists in nature, because the sequence is a compilation of the sequences from various isolates. For example, the synthetic HIV-1 envelope protein described in Example 1, is based on the amino acid sequence of four different HIV 1 isolates, namely, HTLV-IIIB, LAV-1, ARV-2 and CDC-451.

Other properties can be built into the sequence. For example, codons can be switched for optimal expression in bacteria or yeast, specific restriction sites can be introduced, and other restriction sites can be removed. In addition, the sequence should have specific restriction sites at both 5′ and 3′ ends of the fragment to facilitate cloning in a particular vector. Synthetic DNA fragments can be synthesized as follows: (1) select an unique protein sequence, (2) reverse translate to determine complementary DNA sequence, (3) optimize codons for bacterial or yeast expression, and (4) introduce and/or remove specific restriction sites.

Sixty-one distinct nucleotide codons code for 20 amino acids giving rise to the degeneracy in the genetic code. Researchers have reported the frequencies of codons used in nucleic acids for both eukaryotic and prokaryotic organisms. (Grantham et al., *Nucleic Acids Res.* [1980] 9:r43; Gouy et al., *Nucleic Acids Res.* [1982] 10:7055; Sharp et al., *Nucleic Acids Res.* [1986] 14:7737.) Sequences from the entire *E. coli* genome, with examples of sequences from the chromosome, transposons, and plasmids, have been analyzed. These sequences code for structural proteins, enzymes and regulatory proteins. Correlation has been shown between the degree of codon bias within a particular gene and the level of gene expression.

It is preferred that the same codon triplet for each particular amino acid of the synthetic DNA sequence be used. However, alternative codon(s) can be used to add or delete a particular restriction site. The sequence should include unique restriction sites which can be used to delete a specific fragment or sequence, or substitute a particular sequence in case of an error in the original synthesis.

Vector systems which can be used include plant, bacterial, yeast, insect, and mammalian expression systems. It is preferred that the codons are optimized for expression in the system used. The proteins and polypeptides provided by the invention, which are cloned and expressed in heterologous systems, as described above, can be used for diagnostic and therapeutic purposes.

A preferred expression system utilizes the lambda pL vector system. This expression system has the following features: (1) a strong lambda pL promoter, (2) a strong three-frame translation terminator rrnBt1, and (3) translation starts at an ATG codon, eight base pairs from the ribosome binding site located within an accessible NcoI restriction site.

Another advantage of the expression system of the present invention is that one can customize the synthetic DNA fragments so they contain specific DNA sequences which express proteins with desired amino acid sequences, and further allows one the capability of adding, at either the 5′ or 3′ end, other DNA sequences to facilitate the transfer of synthetic fragments into various vectors.

Additionally, the use of particular restriction sites at both ends of the fragment may also facilitate incorporation of the fragment into other sequences to generate fusion proteins, which can also be used as diagnostic and therapeutic reagents. For example, the HIV-1 gp41 sequence can be incorporated within or at the end of core/surface antigen of the hepatitis B viral sequence to generate a fusion protein which can be used in a single assay screening system for the detection of both AIDS and Hepatitis B in prospective blood donors. Alternatively, the assay can be used to track the course of a patient's infection.

Other proteins from any source, including bacterial, yeast, insect, plant or mammalian, can be used with the synthetic DNA fragments of the invention to produce fusion proteins. Those which are expressed efficiently in their respective expression systems are especially preferred because they can enhance the expression of the synthetic fragment of the fusion protein.

The synthetic DNA sequences of the present invention, derived from several HIV isolates and optimized for expression in *E. coli*, provides continuous availability and uniformity of HIV antigen preparations which will recognize test sera from individuals exposed to genetically distinguishable variant HIV isolates. The recombinant antigen expression is very stable since *E. coli* codons have been used for its synthesis. Moreover, the insertion of specific restriction sites allows addition, deletion, or substitution in important antigenic epitopes in the coding sequences, a property difficult to achieve when naturally occurring HIV sequences are utilized for expression. Construction of synthetic genes also allows the addition of protein sequences at either amino- or carboxyl- terminus of the gene thereby allowing the development of novel reagents. For example, a fusion gene can be produced comprising a fusion between HIV-1 core antigen and HIV-1 envelope synthetic gene. More specifically the envelope synthetic gene comprises the carboxyl-terminus HIV-1 gp120 sequence and the full length HIV-1 gp41. Similarly, the HIV-1 core antigen DNA sequence can be fused to the HIV-2 gp41 sequences, both of which can be expressed at high levels in heterologous host systems such as *E. coli*.

E. coli strains containing plasmids useful for constructs of the invention have been deposited at the American Type Culture Collection, Rockville, Maryland, on November 22, 1988, under the accession nos. ATCC 67855 (pSD301/RR1/pRK248.clts) and ATCC 67856 (pSD306/CAG456).

The following examples further describe the invention. The examples are not intended to limit the invention in any manner.

Reagents and enzymes

Media such as Luria-Bertani (LB) and Superbroth II (Dri Form) were obtained from Gibco Laboratories Life Technologies, Inc., Madison, Wisconsin. Restriction enzymes, Klenow fragment of DNA polymerase I, T4 DNA ligase, T4 polynucleotide kinase, nucleic acid molecular weight standards- M13 sequencing system, X-gal (5-bromo-4-chloro-3-indonyl-β-D-galactoside), IPTG (isopropyl-β-D-thiogalactoside), glycerol, Dithiothreitol, 4-chloro-1-napthol were purchased from Boehringer Mannheim Biochemicals, Indianapolis, Indiana; or New England Biolabs, Inc., Beverly, Massachusetts; or Bethesda Research Laboratories Life Technologies, Inc., Gaithersburg, Maryland. Prestained protein molecular weight standards, acrylamide (crystallized, electrophoretic grade >99%); N-N′-Methylene-bis-acrylamide (BIS); N,N,N′,N′,-Tetramethylethylenediamine (TEMED) and sodium dodecylsulfate (SDS) were purchased from BioRad Laboratories, Richmond, California. Lysozyme and ampicillin were obtained from Sigma Chemical Co., St.

Louis, Missouri. Horseradish peroxidase (HRPO) labeled secondary antibodies were obtained from Kirkegaard & Perry Laboratories, Inc., Gaithersburg, Maryland. Seaplaque agarose (low melting agarose) was purchased from FMC Bioproducts, Rockland, Maine.

T50E10 contained 50 mM Tris, pH 8:0, 10 mM EDTA; 1X TG contained 100 mM Tris, pH 7.5 and 10% glycerol; 2X SDS/PAGE loading buffer consisted of 15% glycerol, 5% SDS, 100 mM Tris base, 1M β-mercaptoethanol and 0.8% Bromophenol blue dye; TBS contained 50 mM Tris, pH 8.0, and 150 mM sodium chloride; Blocking solution consisted of 5% Carnation nonfat dry milk in TBS.

Host cell cultures. DNA sources and vectors

E. coli JM103 cells, pUC8, pUC18, pUC19 and M13 cloning vectors were purchased from Pharmacia LKB Biotechnology, Inc., Piscataway, New Jersey; Competent Epicurean™ coli strains XL1-Blue and JM109 were purchased from Stratagene Cloning Systems, La Jolla, California. RR1 cells were obtained from Coli Genetic Stock Center, Yale University, New Haven, Connecticut; and *E. coli* CAG456 cells from Dr. Carol Gross, University of Wisconsin, Madison, Wisconsin. Vector pRK248.clts was obtained from Dr. Donald R. Helinski, University of California, San Diego, California.

General methods

All restriction enzyme digestions were performed according to suppliers' instructions. At least 5 units of enzyme were used per microgram of DNA, and sufficient incubation was allowed to complete digestions of DNA. Standard procedures were used for mini cell lysate DNA preparation, phenol-chloroform extraction, ethanol precipitation of DNA, restriction analysis of DNA on agarose, and low melting agarose gel purification of DNA fragments (Maniatis et al., *Molecular Cloning. A Laboratory Manual* [New York: Cold Spring Harbor, 1982]). Plasmid isolations from *E. coli* strains used the alkali lysis procedure and cesium chloride-ethidium bromide density gradient method (Maniatis et al., supra). Standard buffers were used for T4 DNA ligase and T4 polynucleotide kinase (Maniatis et al., supra).

EXAMPLES

Example 1

Cloning strategy of codon-optimized synthetic HIV-1 envelope protein

In order to develop a synthetic gene encoding the HIV-1 envelope glycoprotein and fragments thereof, the amino acid sequences of four independent HIV-1 viral isolates designated as HTLV-IIIB (BH102), LAV-1 (MAL), ARV-2 (SF), and CDC-451 (CDC42) were compared. A unique amino acid sequence from the four isolates (Fig. 1) was selected to derive a fragment with amino acid residues nos. 552-668 (numbering by Ratner et al., supra). This fragment contained nine amino acid substitutions (8%) as compared to the HTLV-IIIB (BH102) isolate. This amino acid sequence was reverse translated using codons optimized to facilitate high level expression in *E. coli*. The ambiguous nucleotides remaining in the second and/or third base of the codon were assigned to facilitate molecular cloning, and the addition, substitution, or deletion of sequences. The DNA sequence was then subdivided into eight double stranded fragments with unique 6 bp overhangs to direct specific annealing. The sixteen individual oligonucleotides were synthesized on Applied Biosystem 380A DNA synthesizer using methods and reagents recommended by the manufacturer. These purified oligonucleotides were annealed and ligated together to assemble the entire fragment which was digested with BamHI and SalI, ligated into pUC18 and transformed into *E. coli* JM103 cells. A clone designated BS2-10 (Fig. 2) was isolated, restriction mapped and its DNA sequence confirmed using the Sanger dideoxy chain termination method (Sanger et al., *J. Mol. Biol.* (1982) 162:729).

In order to establish that clone BS2-10 expressed this unique HIV-1 transmembrane protein (TMP)

fragment, the BS2-10/JM103 culture was grown at 37° C in 50 ml Luria Broth, in a 250 ml Erlenmeyer flask. When the cultures reached an OD600 of 0.3-0.5, IPTG was added to a final concentration of 1 mM to induce expression. Samples (1.5 ml) were removed at 1 hr intervals, and the cells were pelleted and resuspended to an OD600 of 10.0 in 2X SDS/PAGE loading buffer. Aliquots (15 µl) of the prepared samples were loaded on a 15% SDS/PAGE gel, and the proteins were separated and then electrophoretically transferred to nitrocellulose for immunoblotting. The nitrocellulose sheet containing the transferred proteins was incubated with Blocking Solution for one hour and incubated overnight at 4° C with AIDS patients' sera diluted in TBS containing 5% *E. coli* JM103 lysate. The nitrocellulose sheet was washed three times in TBS, then incubated with HRPO-labeled goat anti-human IgG, diluted in TBS containing 10% fetal calf sera. The nitrocellulose was washed three times with TBS and the color was developed in TBS containing 2 mg/ml 4-chloro-1-napthol, 0.02% hydrogen peroxide and 17% methanol. Clone BS2-10 demonstrated a strongly immunoreactive band with AIDS patients' sera indicating that the synthetic HIV-1 TMP fragment was expressed in *E. coli*. In order to assemble the full length HIV-1 transmembrane protein, as well as the extreme carboxyl-terminal 37 amino acids of gp120, the amino acid sequences of the four isolates described previously were compared to each other to derive a unique amino acid sequence for this gene. After this unique amino acid sequence was reverse translated using codons optimized for *E. coli* expression, the ambiguous nucleotides were assigned as previously described. The full length synthetic HIV-1 envelope gene (FSG) was divided into six additional subfragments. The complete DNA and amino acid sequence of FSG is shown in Fig. 3, indicating the restriction sites and subfragments used for assembly. Fig. 4 is a comparison of the amino acid sequence used to develop the synthetic HIV-1 envelope gene with known amino acid sequences of 13 independent isolates reported in the Los Alamos HIV Data Bank (Meyers et al., *Human Retroviruses and AIDS* (1987), Los Alamos National Laboratory). The Genalign program of Intelligenetics was used to align these sequences, and the alignment demonstrates that FSG (designated SYNGENE in Fig. 4) retains substantial overall sequence homology compared to other known isolates. Alignment parameters and alignment scores of the individual sequences are also shown.

Synthesis and cloning of subfragments

The subfragments located downstream from BS2-10, designated 413-1 through 413-4, were synthesized along with additional sequences containing a BamHI restriction site at the 5′ end and a HindIII restriction site at the 3′ end to facilitate molecular cloning and DNA sequence analysis of the individual subfragments. The subfragments located upstream of BS2-10 were also synthesized with additional sequences containing restriction sites useful for cloning and DNA sequence analysis. The subfragment encoding the carboxyl-terminal gp120 amino acid sequence, designated c-term gp120, contained EcoRI and BamHI restriction sites on the 5′ end and BglII and SmaI restriction sites on the 3′ end. Similarly, subfragment 415 contained a BglII site on the 5′ end and BglII and BamHI restriction sites on the 3′ end. With the exception of the c-term gp120 subfragment, in which both strands were synthesized as described for BS2-10, the remaining subfragments of FSG were synthesized by a method utilizing the Klenow fragment of DNA polymerase I. In this method, oligonucleotides comprising opposite strands of a particular subfragment, which contained ten complementary bases, were synthesized and annealed. The second complementary strand was then filled in by the Klenow fragment of DNA polymerase I in the presence of the four deoxynucleotides in a manner similar to that described by Sanger et al., supra, for DNA sequencing. The resulting double-stranded subfragment was then digested with the appropriate restriction enzymes and cloned into pUC vectors to confirm the DNA sequence, as previously described. Subfragments 413-1 through 413-4 were cloned into pUC18 using the BamHI and HindIII restriction sites common to all. Subfragment c-term gp120 was cloned into pUC8 using the EcoRI and SmaI restriction sites. Subfragment 415 was cloned into the plasmid containing c-term gp120 at the BglII restriction site and screened for proper orientation by restriction mapping. The plasmid DNAs for all subfragments were prepared by the cesium chloride buoyant density gradient method and the individual DNA sequences were confirmed directly from the double-stranded template (Hattori et al., *Nucl. Acid Res.* (1985) 13:7813).

Assembly and cloning of FSG

Subfragments located downstream from BS2-10 were cloned in a stepwise fashion utilizing unique internal restriction sites at the 5′ end and a common HindIII site at the 3′ end. For example, subfragment 413-1 was cloned into BS2-10 at the SalI and HindIII restriction sites to generate clone BS2-10A, into which

413-2 was inserted at the HpaI and HindIII restriction sites to generate clone BS2-108. Similarly, subfragments 413-3 and 413-4 were added using unique EcoRV and SnaBI restriction sites, respectively. The two subfragments located upstream of clone BS2-10, having been cloned together in pUC8, were ligated to BS2-10 as a BamHI fragment. Fig. 5 shows the cloning method used to assemble the synthetic HIV-1 envelope gene in pUC18. The final clone, designated FSG, was restriction mapped to confirm the proper orientation of the BamHI-BamHI fragment.

Example 2

Cloning and expression of FSG in lambda pL Vector Systems

Expression analysis of FSG was carried out in vector systems utilizing the strong lambda pL promoter and the temperature sensitive cI repressor gene (Benard et al., *Gene* (1979) 5:59). The specific vectors used in these analyses are derivatives of pBR322, containing a lambda pL promoter and a synthetic Shine-Dalgarno sequence, followed by restriction sites used for cloning various genes of interest. In addition, these vectors contain the strong three-frame translation terminator rrnBt1. Vector pSDKR816 contains a NcoI restriction site which provides an ATG start codon optimally spaced from the start of transcription. Fig. 6 schematically presents the cloning of FSG into pSDKR816 to generate clone pSD301. Briefly, FSG was digested with HindIII and SmaI, the ends were made blunt by filling in with the Klenow fragment of DNA polymerase I, and the 1209 bp fragment was purified and ligated into pSDKR816 at the NcoI site filled in with the Klenow fragment of DNA polymerase I. After transformation into *E. coli* RR1 cells containing the cIts gene on the compatible vector pRK248, a clone with FSG in the proper orientation was isolated by restriction mapping and designated pSD301. The specific amino acid sequence encoded by pSD301 is presented in Fig. 7 indicating all linker derived sequences (+) and all amino acid substitutions within the HIV-1 envelope sequences not yet identified in any published sequence (*).

Additionally, FSG was cloned as a fusion to the HIV-1 gag protein (amino acid residue nos. 121-407, numbering by Ratner et al., supra) which is highly expressed under control of the lambda pL promoter in vector pKRR955. FSG was digested with AvaI, the ends were made blunt by filling in with the Klenow fragment of DNA polymerase I, and the 1199 bp fragment was purified and ligated into pKRR955 at the SmaI restriction site to form an HIV-1 gag/synthetic env fusion protein (Fig. 6). After transformation into *E. coli* pRK248.cIts/RR1 cells, a clone containing FSG in the proper orientation was identified by restriction mapping and designated pSD302. The specific amino acid sequence of this fusion protein is presented in Fig. 7 indicating all linker derived sequences, HIV-1 gag sequences, and HIV-1 envelope sequences as previously described.

Fifty ml cultures of pSD301 and pSD302 in *E. coli* pRK248.cIts/RR1 cells were grown in Superbroth II media at 30°C to an OD600 of 0.5, at which time the cultures were shifted to 42°C to inactivate the temperature sensitive cI repressor and thereby induce expression off the lambda pL promoter. Two samples (2.0 ml each) were removed at 1 hr intervals. Sample preparation was as follows.

The cells were pelleted, then resuspended in either 1X TG buffer or T50E10 buffer. An equal volume of 2X SDS/PAGE loading buffer was added to the 1X TG suspended cells to produce the whole lysate. The sample resuspended in T50E10 was sonicated eight times for 30 seconds each, at a power setting of 10 watts, using the microtip provided with the Vibra Cell Sonicator (Sonics and Materials, Inc., Danbury, CT). The sonicated sample was then centrifuged to remove the insoluble fraction which was resuspended in the original starting volume of T50E10. An equal volume of 2X SDS/PAGE loading buffer was added to both the sonicated soluble and insoluble fractions, which together with the whole cell lysate, were boiled for 5 min, centrifuged to remove any remaining insoluble material, and aliquots (15µl) were separated on duplicate 12.5% SDS/PAGE gels. Proteins from one such gel were electrophoretically transferred to nitrocellulose for immunoblotting with AIDS patients' sera, as previously described. The second gel was fixed in a solution of 50% methanol, 10% acetic acid for twenty minutes at room temperature, and then stained with 0.25% Coomassie blue dye in a solution of 50% methanol, 10% acetic acid for 30 minutes. Destaining was carried out using a solution of 10% methanol, 7% acetic acid for 3-4 hr, or until a clear background was obtained.

Fig. 8 presents the expression of pSD301 and pSD302 prior to (T0) and four hours post (T4) induction, analyzed by Coomassie blue staining (Fig. 8A) and immunoblotting (Fig. 8B). Samples were pKRR955 (T0 whole cell lysate [lane 1], T4 whole cell lysate [lane 2]); pSD301 (T0 whole cell lysate [lane 3], T4 whole

cell lysate [lane 4], T4 sonicated soluble fraction [lane 5], and T4 sonicated insoluble fraction [lane 6]); and pSD302 (T0 whole cell lysate [lane 7], T4 whole cell lysate [lane 8], T4 sonicated soluble fraction [lane 9], and T4 sonicated insoluble fraction [lane 10]). Molecular weight standards were run in lane 11. Arrows indicate the position of the induced proteins which are clearly visualized in both the whole cell lysate and the sonicated insoluble cell fraction by Coomassie blue staining (Fig. 8A). Lane 2 indicates that pKRR955 expressed the HIV-1 gag protein at a level greater than 25% of total cellular protein, lane 4 indicates that pSD301 expressed the synthetic HIV-1 envelope protein at a level of approximately 12% of total cellular protein, and lane 8 indicates that pSD302 expressed the HIV-1 gag/synthetic env fusion protein at a level of approximately 5% of total cellular protein. The expression levels obtained using FSG were significantly higher than those obtained using the corresponding native viral DNA sequences in similar pL vector systems. All three recombinant proteins were highly reactive with AIDS patients' sera (Fig. 8B). This data demonstrates that the synthetic HIV-1 envelope gene, including the hydrophobic region of the transmembrane protein, can be efficiently expressed in *E. coli*, and the expressed proteins are highly immunoreactive.

Example 3

Synthesis and cloning of synthetic HIV-2 TMP and fragment thereof

The entire HIV-2 transmembrane protein (TMP) was chemically synthesized using the method of oligonucleotide directed double-stranded break repair disclosed in U.S. Patent Application Serial No. 883,242, filed July 8, 1986 by Mandecki (EPO 87109357.1), which is incorporated herein by reference. Envelope amino acid residues 502-858 of the HIV-2 ROD isolate (numbering by Guyader et al., supra) were reverse translated using codon assignments optimal for expression in *E. coli*. After specific nucleotides were assigned to the remaining ambiguous nucleotides, as previously described, the full length TMP sequence was generated as indicated in Fig. 9. The synthetic gene was assembled and cloned as three separate subfragments represented by fragment A, a 335 bp HindIII-NcoI fragment, fragment B, a 309 bp NcoI-BamHI fragment (29 hydrophobic amino acid residues deleted), and fragment C, a 362 bp BamHI-HindIII fragment, as depicted in Fig. 10. A fourth fragment containing the deleted twenty-nine hydrophobic amino acid residues was cloned into the 309 bp NcoI-BamHI fragment as an EcoRV-SnaBI fragment (Fig. 10). The three major subfragments were cloned into pUC vectors, transformed into JM109 cells and their primary nucleotide sequences confirmed, as previously described. The fragments were then gel-purified and ligated together to form the 1089 bp full length synthetic HIV-2 TMP. This 1089 bp HindIII fragment was cloned into pUC8 and designated pJC28 (Figure 11).

Fragment A encoding the amino terminal 108 amino acids of HIV-2 TMP (from Tyr 502 to Trp 609 [Guyader et al., supra]) was cloned at the HindIII-SalI sites of pUC19. A clone, designated pJC22, was identified by restriction mapping and its primary nucleotide sequence was confirmed. Plasmid pJC22 was digested with HindIII-Asp718 to release a 361 bp fragment containing the synthetic HIV-2 TMP gene fragment which was ligated into the HindIII-Asp718 sites of plasmid pTB210 and transformed into *E. coli* XL1 cells. Plasmid pTB210 is disclosed in a U.S. Patent Application entitled "CKS Method of Protein Synthesis", concurrently filed by T. Bolling et al., which is a CIP of an earlier application, U.S. Serial No. 167,067, filed March 11, 1988, which is hereby incorporated by reference. A clone, designated pJC100 (Fig. 12), was isolated and restriction mapped to identify the hybrid gene of kdsB (encoding CKS) and HIV-2 TMP fragment.

Example 4

Cloning of synthetic HIV-2 TMP in lambda pL vectors

The 1089 bp HindIII fragment containing the entire HIV-2 TMP was isolated from pJC28, filled in with

the Klenow fragment of DNA polymerase I to produce blunt ends and cloned directly behind an ATG start codon provided by the filled in NcoI site of pSD305 (pSDKR816 previously described with clts inserted). Similarly, an 1097 bp SalI-Asp718 fragment containing the entire HIV-2 TMP was isolated from pJC28, filled in with the Klenow fragment of DNA polymerase I to produce blunt ends and cloned at the SmaI site of pKRR955 (previously described) to produce an HIV-1 gag/HIV-2 TMP fusion protein. The clone containing the HIV-2 TMP gene under control of the lambda pL promoter was designated pSD306 and the clone containing the HIV-2 TMP as a fusion to HIV-1 gag under control of the lambda pL promoter was designated pSD307, as outlined in Fig. 13. After transformation of pSD306 into *E. coli* CAG456 cells (Baker, *PNAS* (1984) 81:6779) and pSD307 into *E. coli* pRK248.clts/RR1 cells, single cell clones were isolated and restriction mapped to demonstrate the presence and orientation of the HIV-2 TMP gene. The specific amino acid sequences of pSD306 and pSD307 are presented in Fig. 14, indicating linker derived sequences, HIV-1 gag sequences, and HIV-2 TMP sequences. Expression of the synthetic HIV-2 TMP gene was induced in these cultures by temperature shift methods, as previously described. Aliquots of the cultures before and after induction were subjected to SDS/PAGE analysis for both Coomassie blue staining and immunoblotting using HIV-2 positive human sera, as previously described for the synthetic HIV-1 envelope gene product. Whole cell lysates and the sonicated soluble and insoluble fractions of the cultures were analyzed and are illustrated in figures 15 and 16 for the pSD306 and pSD307 constructs, respectively.

Fig. 15 presents the expression of pSD306 prior to (T0) and two hours post (T2) induction, analyzed by Coomassie blue staining (Fig. 15A) and immunoblotting (Fig. 15B). Samples were T0 whole cell lysate (lane 1); T0 sonicated soluble fraction (lane 2); T0 sonicated insoluble fraction (lane 3); T2 whole cell lysate (lane 4); T2 sonicated soluble fraction (lane 5); T2 sonicated insoluble fraction (lane 6); and BioRad prestained molecular weight markers (lane M). Fig. 15 demonstrates that pSD306 expressed a significant amount of the HIV-2 TMP at time T2, as indicated by the arrows on both the Coomassie stained gel and the immunoblot. This expressed protein is visible in both the whole cell lysate as well as the sonicated insoluble cell fraction of these cultures.

Similarly, Fig. 16 presents the expression of pSD307 prior to (T0) and two hours post (T2) induction, analyzed by Coomassie blue staining (Fig. 16A) and immunoblotting (Fig. 16B). Samples were pKRR955, T2 whole cell lysate (lane 1); pSD307, T0 whole cell lysate (lane 2), T0 sonicated soluble fraction (lane 3), T0 sonicated insoluble fraction (lane 4), T2 whole cell lysate (lane 5), T2 sonicated soluble fraction (lane 6), T2 sonicated insoluble fraction (lane 7); and BioRad prestained molecular weight markers (lane M). Fig. 16 demonstrates that pSD307 expressed a significant amount of the HIV-1 gag/HIV-2 TMP fusion protein at time T2, as indicated by the arrows on both the Coomassie stained gel and the immunoblot. This fusion protein is also visible in both the whole cell lysate and the sonicated insoluble fraction of these cultures. The HIV-1 gag fusion partner (lane 1), although present at higher levels than the HIV-1 gag/HIV-2 TMP fusion protein, showed lower immunoreactivity to HIV-2 specific antibodies.

Example 5

Diagnostic utility of synthetic DNA derived HIV proteins

The HIV specific proteins overexpressed in *E. coli* were purified using procedures known in the art. The proteins expressed at high levels were immunogenic and were recognized by antibodies produced in HIV-infected individuals (see figs. 8, 15 and 16). The HIV specific proteins derived from *E. coli* can be utilized in several immunoassay configurations, as described in CIP application U.S. Serial No. 020,282, filed February 27, 1987 by Dawson *et al.*, which is hereby incorporated by reference. The parent application is EPO 86116854.0 (December 4, 1986). In a preferred configuration, HIV specific proteins were coated on solid support and incubated with test samples. The virus specific antibodies present in the test sample recognized and were bound to the HIV proteins on the solid support. The HIV specific antibodies were quantitated by the use of goat anti-human immunoglobulin conjugated to HRPO.

The HIV-1 exposed individuals were detected by the use of HIV-1 specific proteins, such as HIV-1 gp41 and HIV-1 p24 proteins derived by recombinant DNA techniques, described in the CIP application Serial No. 020,282. However, only 70 to 90% of the HIV-2 exposed individuals were detected using these HIV-1 specific proteins, due to cross reactivity between the two strains. The HIV-2 exposed individuals which were not detected using these HIV-1 specific proteins were detected using synthetic DNA derived HIV-2 proteins.

For example, the HIV-2 TMP fragment fused to CKS (pJC100) when supplemented to the recombinant HIV-1 proteins on the solid support described above significantly increased the detection of test samples containing HIV-2 antibodies as illustrated in Table 1, below.

Table 1

	HIV-1 Test	HIV-1/HIV-2 Test
Samples Tested*	127	127
Non Reactive	26 (20.47%)	0 (0%)
Reactive	101 (79.53%)	127 (100%)

* All 127 samples were confirmed positive for the presence of HIV-2 antibodies by western blot analysis using disrupted HIV-2 virus.

Additionally, 3,411 normal blood donors were screened using the HIV-1/HIV-2 recombinant assay described above. The recombinant assay demonstrated a specificity of 99.77%, with only eight (0.23%) initial reactive and four (0.12%) repeat reactive samples.

Example 6

Differentiation of HIV-1 and HIV-2 infections

Frequently, individuals who have been exposed to HIV-2 have antibodies directed against epitopes on HIV-2 proteins which are also present on HIV-1 proteins. Likewise, individuals who have been exposed to HIV-1 have antibodies which cross-react with HIV-2 proteins. Because most of the cross-reactions are related to the gag gene products, the pJP100 protein and a recombinant protein from HIV-1 envelope protein (described in CIP Application Serial Number 020,282) were used to differentiate between individuals infected with HIV-1 and HIV-2.

Two independent enzyme-linked immunoassays were developed. Test 1 used HIV-1 recombinant proteins coated upon a solid phase. Test 2 used HIV-2 TMP (pJP100) coated upon a solid phase. Specimens from HIV seropositive individuals from the United States, Portugal or West Africa were tested for antibodies using these two tests. Endpoint titers were determined by diluting the specimens in normal human plasma and testing the dilutions. As illustrated in Table 2 below, specific tests using synthetic recombinant proteins can be effectively used to differentiate HIV-1 and HIV-2 infections.

Table 2

Specimen	Test 1 Endpoint Titer	Test 2 Endpoint Titer
Chicago-AIDS-1	256	<1
Chicago-AIDS-2	512	<1
Chicago-AIDS-3	512	<1
Chicago-Asymptomatic-4	1024	<1
Chicago-Asymptomatic-5	2048	<1
Chicago-Asymptomatic-6	512	<1
West Africa-1	<1	2048
West Africa-2	<1	64
Portugal-1	<1	512

Biological samples which are easily tested by the methods of the present invention include human and animal body fluids such as whole blood, serum, plasma, urine, saliva, stools, lymphocyte or cell culture preparations and purified and partially purified immunoglobulins. The polypeptides and fragments described herein can be used to determine the presence or absence of antibodies to HIV-1 and HIV-2 antigens by assay methods known to those skilled in the art, and for distinguishing between HIV-1 and HIV-2 infections.

One such assay involves:

a) coating a solid support with the polypeptides and polypeptide fragments disclosed herein;

b) contacting the coated solid support with the biological sample to form an antibody polypeptide complex;

c) removing unbound biological sample from the solid support;

d) contacting the complex on the solid support with a labeled immunoglobulin specific for the antibody; and

e) detecting the label to determine the presence or absence of HIV-1 and/or HIV-2 antibodies in the biological sample.

A second assay method involves:

a) coating a solid support with the polypeptides and polypeptide fragments disclosed herein;

b) contacting the coated solid support with the biological sample and the homologous polypeptides conjugated to a label;

c) removing unbound biological sample and unbound labeled polypeptide; and

d) detecting the label to determine the presence or absence of HIV-1 and/or HIV-2 antibodies in the biological sample.

Solid supports which can be used in such immunoassays include wells of reaction trays, test tubes, beads, strips, membranes, filters, microparticles or other solid supports which are well known to those skilled in the art. Enzymatic, radioisotopic, fluorescent, chemiluminescent and colloidal particle labels can be used in the aforementioned assays. Furthermore, hapten/labeled anti-hapten systems such as a biotin/labeled anti-biotin system can be utilized in the inventive assays. Both polyclonal and monoclonal antibodies are useful as reagents, and IgG as well as IgM class HIV antibodies may be used as solid support or labeled reagents.

It is evident from the foregoing examples that one skilled in the art could clone together specific subfragments of the synthetic genes constructed to generate new synthetic genes that would have the same characteristics as those illustrated herein. For example, the c-term gp120 subfragment, BS2-10 and subfragment 413-1 can be cloned together to produce synthetic gene products useful as diagnostic and therapeutic reagents for AIDS.

Claims

1. A polypeptide comprising an amino acid sequence represented by the following:
MGDPMMRDNWRSELYKYKVVKIEPLGIAPTKAKRRVVQREKRADLAVGILGALFLGFLGAAGSTMGARSL
TLTVQARQLLSGIVQQQNNLLRAIKDPKAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKL ICT-
TAVPWNASWSNKSLEDIWNNMTWMQWEREINNYTNLIYSLLEESQNQQEKNEQELLQLDKWVDASLW

NWSNITKWLWYIKLFIMIVGGLAGLRIVFAVLSIVNRVRQGYSPLSFQTRLPNPRGPDRPEGIDEEGGER
DRDRSTRLVDISLALVWEDLRSLCLFSYHRLRDLLLIATRIVELLGRRGWEVLKYWWNLLQYVSQELKNS AVSL-
VNATAIAVAEGTDRVIEVVQRAYRAIRHIHRRIRQGLERILLQVHASSLESSWQFGPG.

2. A polypeptide comprising an amino acid sequence represented by the following:

MSLKIYSSAHGRHTRGVFVLGFLGFLATAGSAMGAASLTVSAQSRTLLAGIVQQQQQLLDVVKRQQELLR
LTVWGTKNLQARVTAIEKYLQDQARLNSWGCAFRQVCHTTVPWVNDSLAPDWDNMTWQEWEKQVRYLEAN
ISKSLEQAQIQQEKNMYELQKLNSWDIFGNWFDLTSWVKYIQYGVLIIVAVIALRIVIYVVQMLSRLRKG YRPVFS-
SPPGYIQQIHIHKDRGQPANEETEEDGGSNGGDRYWPWPIAYIHFLIRQLIRLLTRLYSICRDL
LSRSFLTLQLIYQNLRDWLRLRTAFLQYGCEWIQEAFQAAAARATRETLAGACRGLWRVLERIGRGILAVP RRIR-
QGAEIALLVPSSWQFGPG.

3. A polypeptide fragment of the polypeptide of Claim 2 comprising an amino acid sequence represented by the following:

YSSAHGRHTRGVFVLGFLGFLATAGSAMGAASLTVSAQSRTLLAGIVQQQQQLLDVVKRQQELLRLTVWG
TKNLQARVTAIEKYLQDQARLNSWGCAFRQVCHTTVPW

4. The polypeptide of Claims 1, 2, or 3 produced by E. coli.

5. A fusion polypeptide comprising a polypeptide as in one of Claims 1-3, in which the polypeptide is fused to a prokaryotic or eukaryotic protein.

6. The fusion polypeptide of Claim 5 wherein said prokaryotic or eukaryotic protein is the E. coli enzyme CKS.

7. A synthetic gene comprising a DNA sequence represented by the following:

ATGGGGGATCCCATGATGCGCGACAACTGGCGCTCTGAACTGTACAAATACAAAGTTGTTAAAATCGAAC
CGCTGGGCATCGCTCCGACCAAAGCTAAACGCCGCGTTGTTCAGCGCGAAAAACGCGCAGATCTAGCTGT
TGGTATCCTGGGTGCTCTGTTTCTGGGTTTTCTGGGTGCTGCTGGTTCTACTATGGGTGCTCGCTCTCTG
ACTCTGACTGTTCAGGCTCGCCAGCTGCTGTCTGGTATCGTTCAGCAGCAGAACAACCTGCTGCGCGCTA
TCAAGGATCCCAAAGCTCAGCAGCATCTGCTGCAACTGACTGTTTGGGGTATCAAACAACTGCAGGCTCG
CGTTCTGGCTGTTGAACGCTACCTGAAAGACCAGCAGCTGCTGGGTATCTGGGGTTGCTCTGGTAAACTG
ATTTGCACTACTGCCGTTCCGTGGAACGCTTCTTGGTCCAACAAATCTCTGGAAGACATCTGGAACAACA
TGACTTGGATGCAATGGGAACGCGAAATCAACAACTACACTAACCTGATCTACTCTCTGCTGGAAGAATC
TCAGAACCAGCAGGAAAAAAACGAACAGGAACTGCTGCAACTGGACAAATGGGTCGACGCTTCTCTGTGG
AACTGGTCTAACATAACTAAATGGCTGTGGTACATCAAACTGTTTATCATGATCGTTGGTGGTCTGGCCG
GCCTGCGCATCGTTTTTGCTGTTCTGTCTATCGTTAACCGCGTTCGCCAGGGTTACTCTCCGCTGTCTTT
TCAGACTCGCCTGCCGAACCCGCGCGGTCCGGACCGCCCGGAAGGTATCGATGAAGAAGGTGGTGAACG-
C
GACCGCGACCGCTCTACTCGCCTGGTAGATATCTCTCTGGCTCTGGTTTGGGAAGACCTGCGCTCTCTGT
GCCTGTTTTCTTACCATCGCCTGCGCGACCTGCTGCTGATCGCTACTCGCATCGTTGAACTGCTGGGTCG
CCGCGGTTGGGAAGTGCTGAAATACTGGTGGAACCTGCTGCAATACGTATCTCAGGAACTGAAAAACTCT
GCTGTTTCTCTGGTTAATGCTACTGCTATCGCTGTTGCTGAAGGTACTGACCGCGTTATCGAAGTTGTTC
AGCGCGCTTACCGCGCTATCCGCCATATCCATCGCCGCATCCGCCAGGGTCTGGAACGCATCCTGCTGCA
GGTGCATGCCTCGAGTCTAGAAAGCTCATGGCAATTCGGGCCCGGGTAA

8. The synthetic gene of Claim 7 coding for the polypeptide of Claim 1.

9. A synthetic gene comprising a DNA sequence represented by the following:

ATGAGCTTAAAGATCTACTCTTCCGCTCACGGCCGTCACACCCGTGGCGTTTTCGTTCTGGGCTTCCTGG
GCTTCCTGGCTACCGCGGGCTCCGCTATGGGCGCTGCTTCCCTGACCGTTTCCGCTCAGTCCCGTACCCT
GCTGGCTGGCATCGTTCAGCAGCAGCAGCAACTTCTAGACGTTGTTAAACGTCAGCAGGAGCTCCTGCGT
CTGACCGTTTGGGGCACCAAAAACCTGCAGGCTCGTGTTACCGCTATCGAAAAATACCTGCAGGACCAGG
CTCGTCTGAATTCCTGGGGCTGCGCTTTCCGTCAGGTTTGCCACACCACCGTTCCATGGGTTAACGATTC
CCTGGCTCCGGACTGGGACAACATGACCTGGCAGGAATGGGAAAAACAGGTTCGTTACCTGGAAGCTAAC
ATCTCCAAATCCCTGGAACAGGCTCAGATCCAGCAGGAAAAAAACATGTACGAACTGCAGAAACTGAACT
CCTGGGATATCTTCGGCAACTGGTTCGACCTGACCTCCTGGGTTAAATATATCCAGTACGGCGTGCTCAT
CATCGTTGCTGTTATCGCTCTGCGTATCGTTATCTACGTAGTTCAGATGCTGTCCCGTCTGCGTAAAGGC
TACCGTCCGGTTTTCTCTTCCCCCCCGGGCTATATCCAGCAGATCCATATCCACAAAGACCGTGGCCAGC
CGGCTAACGAAGAAACCGAAGAAGACGGCGGATCCAACGGCGGCGACCGTTACTGGCCGTGGCCGATCG-
C TTATATCCACTTCCTGATCCGTCAGCTGATCCGTCTGCTGACCCGTCTATACTCCATCTGCCGTGACCTG
CTGTCCCGTTCCTTCCTGACCCTGCAACTGATCTACCAGAACCTGCGTGACTGGCTGCGTCTGCGTACCG
CTTTCCTGCAGTACGGCTGCGAATGGATTCAGGAAGCATTCCAAGCGGCCGCTCGTGCTACCCGTGAAAC
CCTGGCTGGCGCATGCCGTGGCCTGTGGCGTGTTCTGGAACGTATCGGCCGTGGTATCCTGGCTGTTCC-
G
CGTCGTATCCGTCAGGGCGCCGAAATCGCTCTGCTGGTACCAAGCTCATGGCAATTCGGGCCCGGGTAA

10. The synthetic gene of Claim 9 coding for the polypeptide of Claim 2.

11. A synthetic gene comprising a DNA sequence represented by the following: TACTCTTCCGCTCACGGCCGTCACACCCGTGGCGTTTTCGTTCTGGGCTTCCTGGGCTTCCTGGCTACCG CGGGCTCCGCTATGGGCGCTGCTTCCCTGACCGTTTCCGCTCAGTCCCGTACCCTGCTGGCTGGCATCGT TCAGCAGCAGCAGCAACTTCTAGACGTTGTTAAACGTCAGCAGGAGCTCCTGCGTCTGACCGTTTGGGGC ACCAAAAACCTGCAGGCTCGTGTTACCGCTATCGAAAAATACCTGCAGGACCAGGCTCGTCTGAATTCCT GGGGCTGCGCTTTCCGTCAGGTTTGCCACACCACCGTTCCATGG

12. The synthetic gene of Claim 11 coding for the polypeptide of Claim 3.

13. A fusion polypeptide comprising a polypeptide as in one of Claims 1-3, in which the polypeptide is fused to a HIV gag protein.

14. The fusion polypeptide of Claim 13 wherein said HIV gag protein is an HIV-1 gag protein comprising an amino acid sequence represented by the following: DTGHSSQVSQNYPIVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNT VGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGE IYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDC KTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNTATIMMQRGNFRNQRKMVKCFNCGKEGH TARNCRA

15. A method for detecting antibodies to HIV antigens in an individual which comprises the steps of:

a) obtaining a sample of a body fluid from the individual:

b) incubating said body fluid with said polypeptide of Claims 1, 2 or 3;

c) incubating said body fluid with a labeled antibody to immunoglobulin; and

d) detecting said label and determining therefrom the presence or absence of antibodies to HIV antigens.

16. A method for detecting antibodies to HIV antigens in an individual which comprises the steps of:

a) obtaining a sample of a body fluid from the individual:

b) incubating said body fluid with said polypeptide of Claims 1, 2 or 3;

c) incubating said body fluid with a labeled antigen; and

d) detecting said label determining therefrom the presence or absence of antibodies to HIV antigens.

(7)

Clustered order of selected sequences:

2.	CDC42FRAG.PEP	(1-107)
3.	BH102FRAG.PEP	(1-107)
4.	SF2FRAG.PEP	(1-107)
1.	MALFRAG.PEP	(1-107)
5.	SYNFRAG.PEP	(1-107)

```
2     1 kAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGfWGCSGKLICTTAVPWNASWSNKtLdQIWNNMT
3     1 EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWNNMT
1     1 EAQQHLLQLTVWGIKQLQARVLAVERYLrDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEDIWdNMT
5     1 EAQQHLLQLTVWGIKQLQARVLAVERYLqDQrLLGmWGCSGKhICTTfVPWNsSWSNrSLdDIWnNMT
      1 KAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEDIWNNMT

2    69 WMEWDREIdNYThLIytLIEESQNQQEKNqQELLqLDKW
3    69 WMEWDREInNYTsLIhsLIEESQNQQEKNEQELLELDKW
4    69 WMQWEREIdNYTntIYtLIEESQNQQEKNEQELLELDKW
1    69 WMQWEkEIsNYTgiIYnLiEESQiQQEKNEkELLELDKW
5    69 WMQWEREINNYTNLIYSLLEESQNQQEKNEQELLQLDKW
```

FIGURE 1

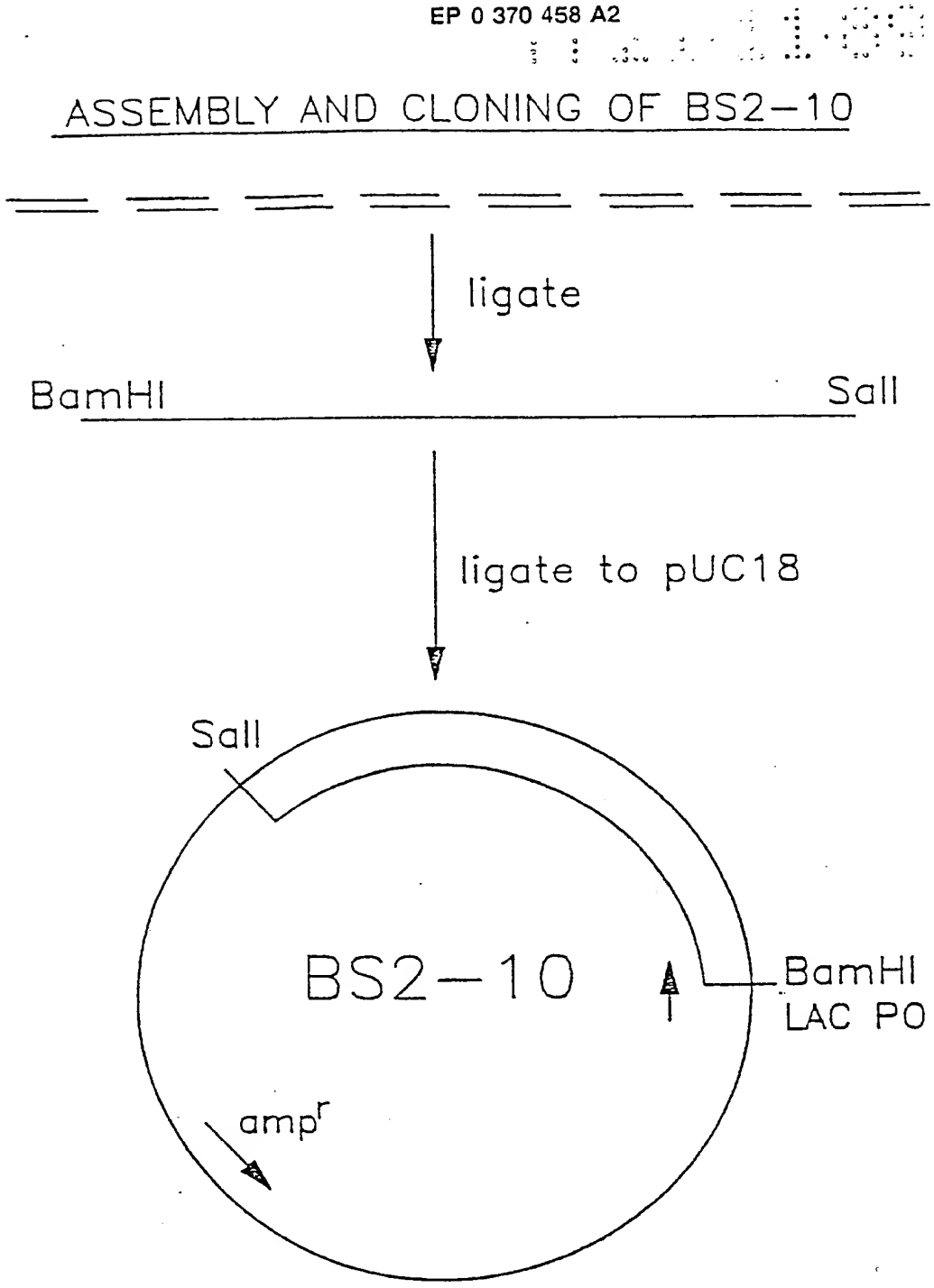

ASSEMBLY AND CLONING OF BS2–10
ligate
BamHI
SalI
ligate to pUC18
SalI
BS2–10
BamHI
LAC PO
amp^r

FIGURE 2

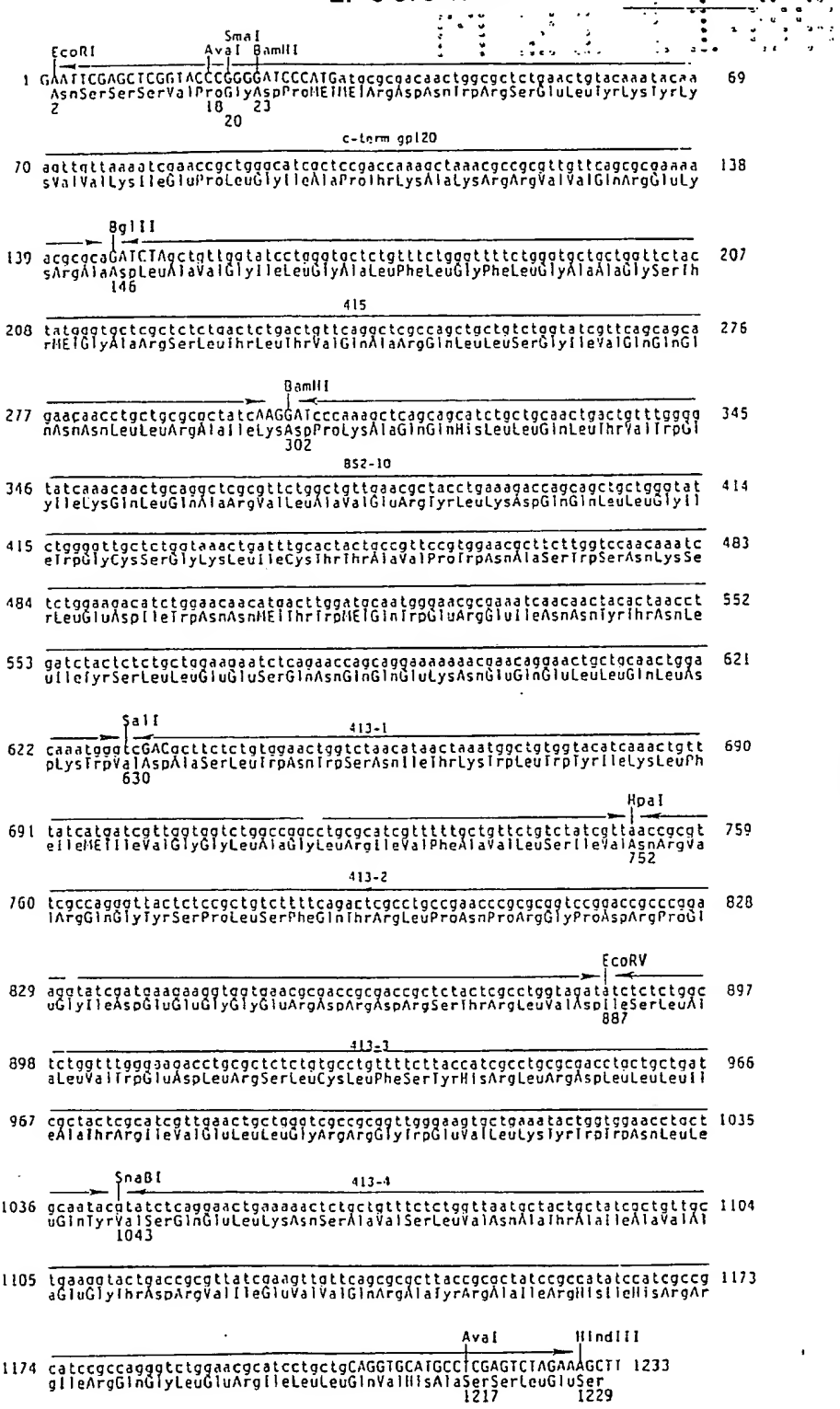


FIGURE 3

```
Amino Alphabet      = Identity
Output line length  = 80
Compress            = Off
Randomization       = Off

AMINO-Res-length    = 2
DELetion-weight     = 1.00
LEngth-factor       = 0
Matching-weight     = 1.00
NUCLEIC-Res-length  = 4
SPread-factor       = 50
```

FIGURE 4-1

Clustered order of selected sequences:

9.	MAL	(1-384)
10.	ELI	(1-383)
13.	Z6	(1-383)
4.	CDC42	(1-384)
12.	RF	(1-384)
11.	WMJ22	(1-384)
7.	BH8	(1-383)
8.	PV22	(1-383)
2.	BRU	(1-383)
1.	HXB2	(1-383)
6.	BH102	(1-383)
14.	HXB3	(1-383)
3.	SF2	(1-384)
5.	SYNGENE	(1-413)

Region Alignment: (listed in Clustered order)

```
 9    1       MRDNWiSELYKYKVVrIEPLGVAPTkAKRRVVEREKRA     IGLGAMFLGFLGAAGSTMGA
10    1       MRDNWRSELYKYKVVqIEPLGVAPTrAKRRVVEREKRA     IGLGAMFLGFLGAAGSTMGA
13    1       MRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVEREKRA     IGLGAMFLGFLGAAGSTMGA
 4    1       MRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRA   VG mLGAMFLGFLGAAGSTMGA
12    1       MRDNWRSELYKYKVVRIEPLGVAPTrAKRRVVQREKRA   VGTI GAMFLGFLGAAGSTMGA
11    1       MRDNWRSELYKYKVVRIEPLGVAPTKAKRRVVQREKRA   VGTI GAMFLGFLGAAGSTMGA
 7    1       MRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRA   VG I GALFLGFLGAAGSTMGA
 8    1       MRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRA   VG I GALFLGFLGAAGSTMGA
 2    1       MRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRA   VG I GALFLGFLGAAGSTMGA
 1    1       MRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRA   VG I GALFLGFLGAAGSTMGA
 6    1       MRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRA   VG I GALFLGFLGAAGSTMGA
14    1       MRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRA   VG I GALFLGFLGAAGSTMGA
 3    1       MRDNWRSELYKYKViKIEPLGIAPTKAKRRVVQREKRA   VG IvGAmFLGFLGAAGSTMGA
 5    1 mgdpmMRDNWRSELYKYKVvKIEPLGIAPTKAKRRVVQREKRAdlaVG IlGAlFLGFLGAAGSTMGA
                                                  +++
```

FIGURE 4-2

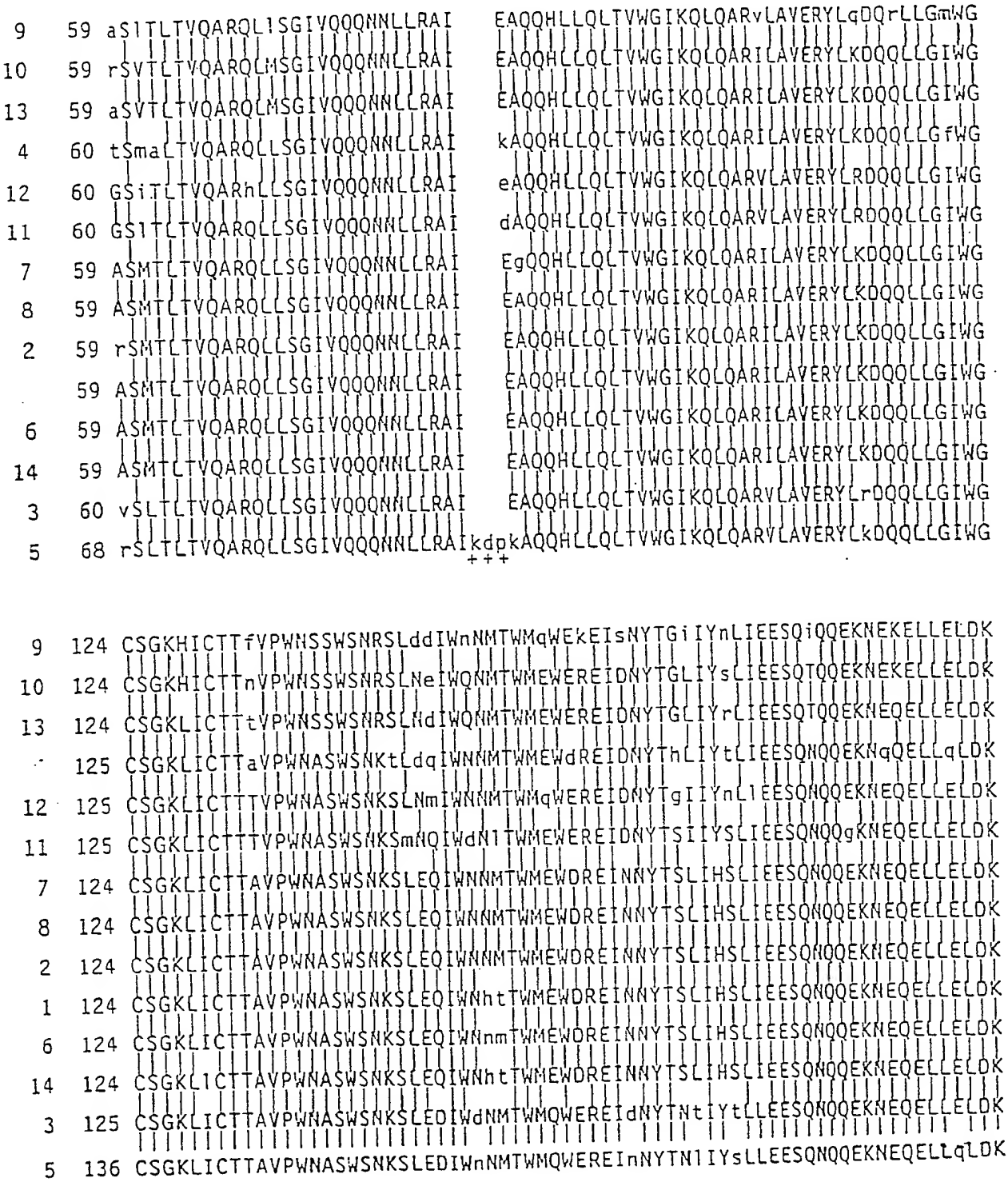

9 59 aS1TLTVQARQL1SGIVQQQNNLLRAI EAQQHLLQLTVWGIKQLQARvLAVERYLqDQrLLGmWG
10 59 rSVTLTVQARQLMSGIVQQQNNLLRAI EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWG
13 59 aSVTLTVQARQLMSGIVQQQNNLLRAI EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWG
4 60 tSmaLTVQARQLLSGIVQQQNNLLRAI kAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGfWG
12 60 GSiTLTVQARhLLSGIVQQQNNLLRAI eAQQHLLQLTVWGIKQLQARVLAVERYLRDQQLLGIWG
11 60 GS1TLTVQARQLLSGIVQQQNNLLRAI dAQQHLLQLTVWGIKQLQARVLAVERYLRDQQLLGIWG
7 59 ASMTLTVQARQLLSGIVQQQNNLLRAI EgQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWG
8 59 ASMTLTVQARQLLSGIVQQQNNLLRAI EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWG
2 59 rSMTLTVQARQLLSGIVQQQNNLLRAI EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWG
59 ASMTLTVQARQLLSGIVQQQNNLLRAI EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWG
6 59 ASMTLTVQARQLLSGIVQQQNNLLRAI EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWG
14 59 ASMTLTVQARQLLSGIVQQQNNLLRAI EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWG
3 60 vSLTLTVQARQLLSGIVQQQNNLLRAI EAQQHLLQLTVWGIKQLQARVLAVERYLrDQQLLGIWG
5 68 rSLTLTVQARQLLSGIVQQQNNLLRAIkdpkAQQHLLQLTVWGIKQLQARVLAVERYLkDQQLLGIWG
9 124 CSGKHICTTfVPWNSSWSNRSLddIWnNMTWMqWEkEIsNYTGiIYnLIEESQiQQEKNEKELLELDK
10 124 CSGKHICTTnVPWNSSWSNRSLNeIWQNMTWMEWEREIDNYTGLIYsLIEESQTQQEKNEKELLELDK
13 124 CSGKLICTTtVPWNSSWSNRSLNdIWQNMTWMEWEREIDNYTGLIYrLIEESQTQQEKNEQELLELDK
125 CSGKLICTTaVPWNASWSNKtLdqIWNNMTWMEWdREIDNYThLIYtLIEESQNQQEKNqQELLqLDK
12 125 CSGKLICTTTVPWNASWSNKSLNmIWNNMTWMqWEREIDNYTgIIYnL1EESQNQQEKNEQELLELDK
11 125 CSGKLICTTTVPWNASWSNKSmNQIWdN1TWMEWEREIDNYTSIIYSLIEESQNQQgKNEQELLELDK
7 124 CSGKLICTTAVPWNASWSNKSLEQIWNNMTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDK
8 124 CSGKLICTTAVPWNASWSNKSLEQIWNNMTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDK
2 124 CSGKLICTTAVPWNASWSNKSLEQIWNNMTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDK
1 124 CSGKLICTTAVPWNASWSNKSLEQIWNhtTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDK
6 124 CSGKLICTTAVPWNASWSNKSLEQIWNnmTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDK
14 124 CSGKLICTTAVPWNASWSNKSLEQIWNhtTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDK
3 125 CSGKLICTTAVPWNASWSNKSLEDIWdNMTWMQWEREIdNYTNtIYtLLEESQNQQEKNEQELLELDK
5 136 CSGKLICTTAVPWNASWSNKSLEDIWnNMTWMQWEREInNYTN1IYsLLEESQNQQEKNEQELLqLDK

FIGURE 4-3

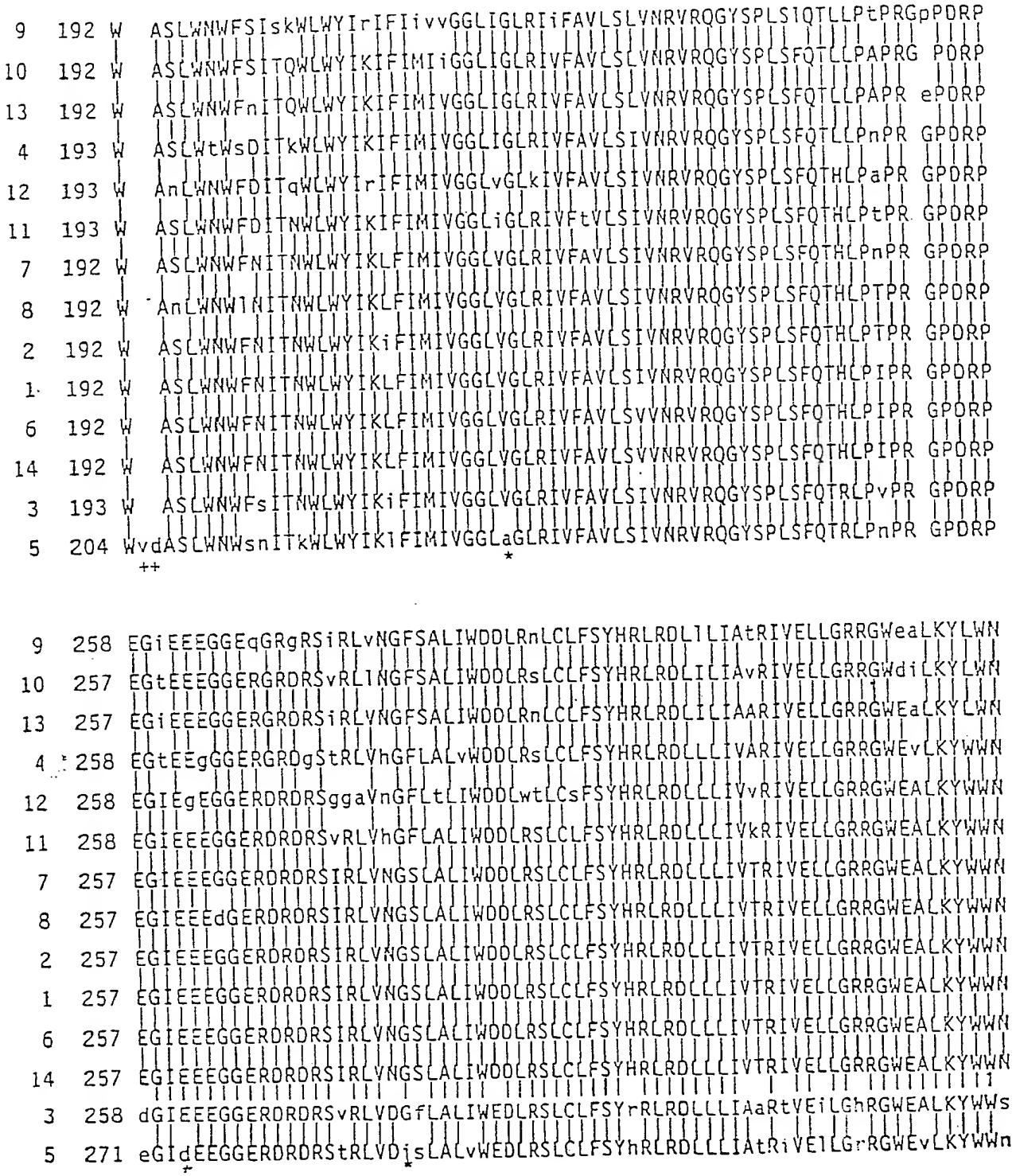


FIGURE 4-4

```
 9  326 LLQYWgQELkNSAiSLlnttAIAVAEcTDRVIEIgQRfgRAiLhIPRRIRQGfERaLL.
10  325 LLQYWSQELRNSASSLfDaIAIAVAEGTDRVIEIiQRacRAVLNIPRRIRQGLERsLL.
13  325 LLQYWSrELRNSASSLlDtIAIAVAEGTDRVIEIVrRtYRAVLNvPtRIRQGLERlLL.
 4  326 LLQYWSQELKNSAVSLvNvTAIAVAEGTDRVIEVVQRIYRAFLHIPRRIRQGfERALL.
12  326 LLQYWSQELKNSAVSLLNtTAIAVAEGTDRiIEVaQRIlRAFLHIPRRIRQGLERALL.
11  326 LLQYWSkELKNSAVgLLNAiAIAVAEGTDRVIEVVQRIcRAIiHIPRRIRQGLERALL.
 7  325 LLQYWSQELKNSAVnLLNATAIAVAEGTDRVIElVQaAYRAIRHIPRRIRQGLERILL.
 8  325 LLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGAYRAIRHIPRRIRQGLERILL.
 2  325 LLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERILL.
 1  325 LLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERILL.
 6  325 LLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGAYRAIRHIPRRIRQGLERILL.
14  325 LLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQeAYRAIRHIPRRIRQGLERILL.
 3  326 LLQYWiQELKNSAVSwLNATAIAVtEGTDRVIEVaQRAYRAIlHIHRRIRQGLERlLL.
 5  339 LLQYvsQELKNSAVSlvNATAIAVaEGTDRVIEVvQRAYRAIrHIHRRIRQGLERiLLqvhassless
          *

 9  385
10  384
13  384
 4  385
12  385
11  385
 7  384
 8  384
 2  384
 1  384
 6  384
14  384
 3  385
 5  407 wqfgpg.
```

Alignment score = 4393.00

Scoring matrix:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1		378	338	339	320	379	375	376	320	327	348	335	331	379
2			341	342	322	378	375	377	323	331	351	338	333	374
3				328	316	340	338	338	315	317	339	333	321	337
4					308	341	340	340	311	322	338	329	325	338
5						322	321	320	284	288	305	296	290	319
6							377	378	322	328	347	336	334	379
7								375	320	326	346	334	332	374
8									320	325	346	336	331	374
9										328	320	320	332	319
10											326	320	355	325
11												345	328	345
12													323	333
13														331
14														

FIGURE 4-5

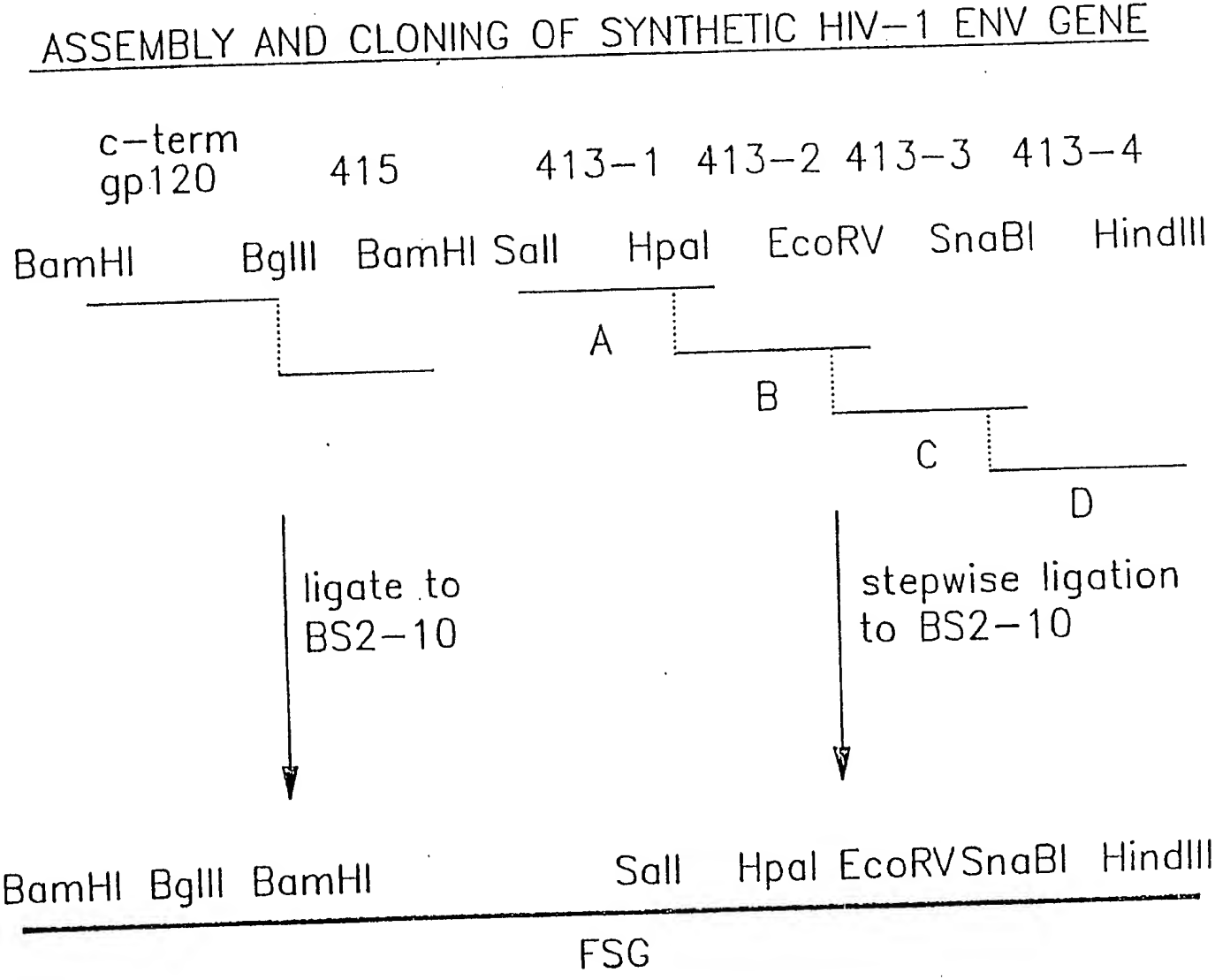


FIGURE 5

EP 0 370 458 A2

EXPRESSION OF SYNTHETIC HIV-1 ENVELOPE

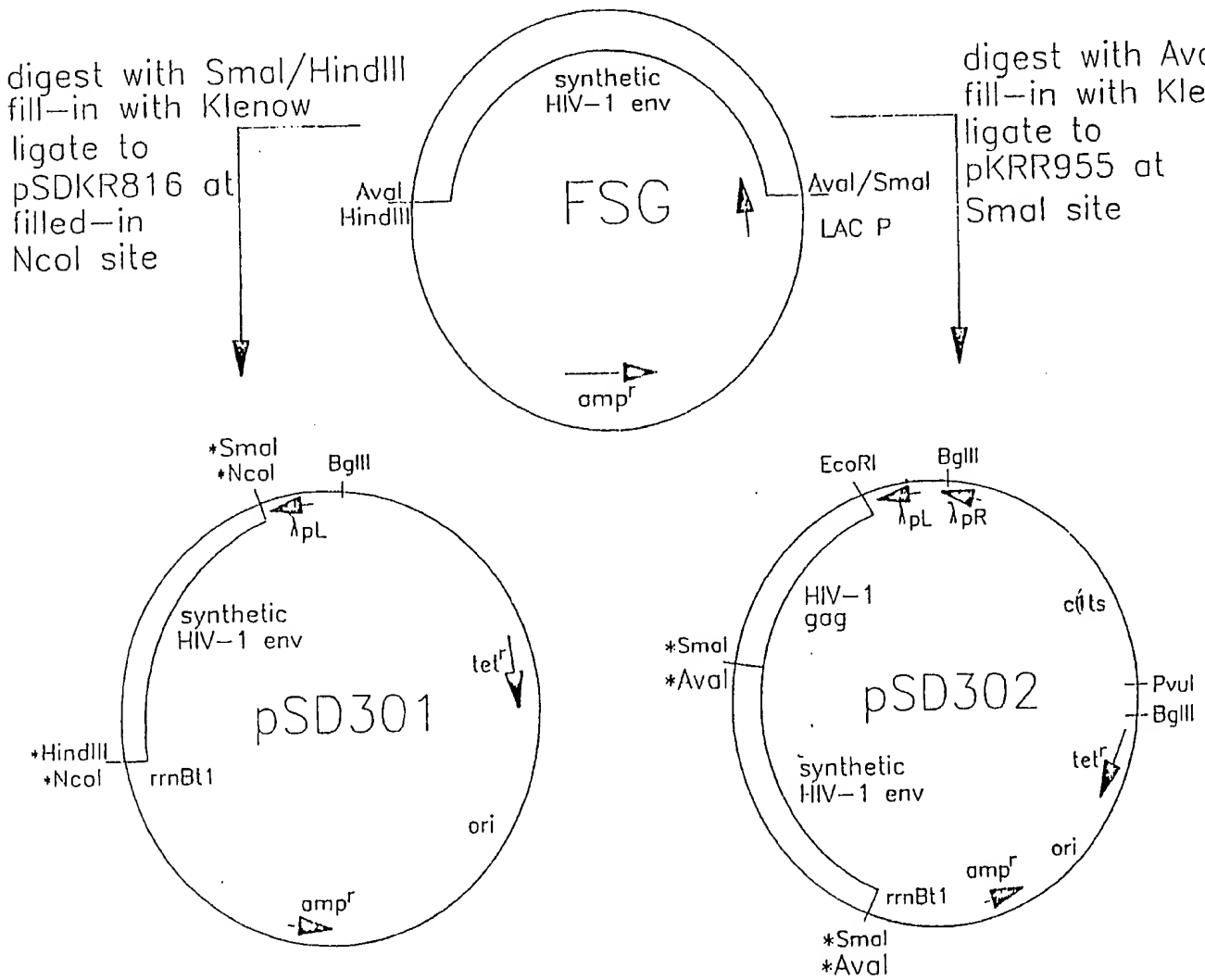


FIGURE 6

PSD301.PEP

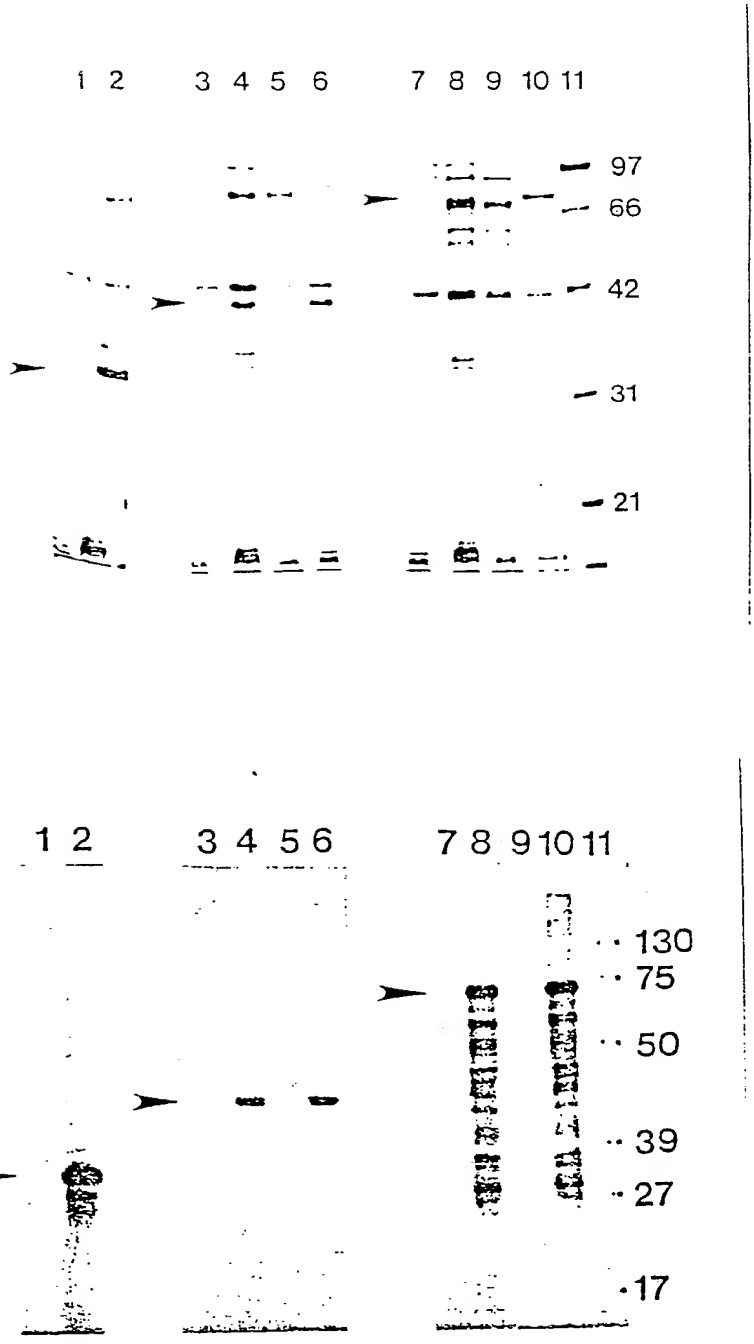
```
          10         20         30         40         50         60         70
MGDPMMRDNW RSELYKYKVV KIEPLGIAPT KAKRRVVQRE KRADLAVGIL GALFLGFLGA AGSTMGARSL
                                                    +++
linker├ HIV-1 env seq ──→
          80         90        100        110        120        130        140
TLTVQARQLL SGIVQQQNNL LRAIKDPKAQ QHLLQLTVWG IKQLQARVLA VERYLKDQQL LGIWGCSGKL
                    +++
         150        160        170        180        190        200        210
ICTTAVPWNA SWSNKSLEDI WNNMTWMQWE REINNYTNLI YSLLEESQNQ QEKNEQELLQ LDKWVDASLW
                                                                   ++
         220        230        240        250        260        270        280
NWSNITKWLW YIKLFIMIVG GLAGLRIVFA VLSIVNRVRQ GYSPLSFQTR LPNPRGPDRP EGIDEEGGER
                     *                                                  *
         290        300        310        320        330        340        350
DRDRSTRLVD ISLALVWEDL RSLCLFSYHR LRDLLLIATR IVELLGRRGW EVLKYWWNLL QYVSQELKNS
          *                                                             *
         360        370        380        390        400        410
AVSLVNATAI AVAEGTDRVI EVVQRAYRAI RHIHRRIRQG LERILLQVHA SSLESSWQFG PG.
                                                       ├ linker seq ──────
```

PSD302.PEP

```
          10         20         30         40         50         60         70
MTMITPSLAA GPDTGHSSQV SQNYPIVQNI QGQMVHQAIS PRTLNAWVKV VEEKAFSPEV IPMFSALSEG
linker seq ──────┼HIV-1 gag seq ──→
          80         90        100        110        120        130        140
ATPQDLNTML NTVGGHQAAM QMLKETINEE AAEWDRVHPV HAGPIAPGQM REPRGSDIAG TTSTLQEQIG
         150        160        170        180        190        200        210
WMTNNPPIPV GEIYKRWIIL GLNKIVRMYS PTSILDIRQG PKEPFRDYVD RFYKTLRAEQ ASQEVKNWMT
         220        230        240        250        260        270        280
ETLLVQNANP DCKTILKALG PAATLEEMMT ACQGVGGPGH KARVLAEAMS QVTNTATIMM QRGNFRNQRK
         290        300        310        320        330        340        350
MVKCFNCGKE GHTARNCRAP GDPMMRDNWR SELYKYKVVK IEPLGIAPTK AKRRVVQREK RADLAVGILG
                     ├linker├ HIV-1 env seq ──→                        +++
         360        370        380        390        400        410        420
ALFLGFLGAA GSTMGARSLT LTVQARQLLS GIVQQQNNLL RAIKDPKAQQ HLLQLTVWGI KQLQARVLAV
                                                    +++
         430        440        450        460        470        480        490
ERYLKDQQLL GIWGCSGKLI CTTAVPWNAS WSNKSLEDIW NNMTWMQWER EINNYTNLIY SLLEESQNQQ
         500        510        520        530        540        550        560
EKNEQELLQL DKWVDASLWN WSNITKWLWY IKLFIMIVGG LAGLRIVFAV LSIVNRVRQG YSPLSFQTRL
                    ++                                  *
         570        580        590        600        610        620        630
PNPRGPDRPE GIDEEGGERD RDRSTRLVDI SLALVWEDLR SLCLFSYHRL RDLLLIATRI VELLGRRGWE
                     *                   *
         640        650        660        670        680        690        700
VLKYWWNLLQ YVSQELKNSA VSLVNATAIA VAEGTDRVIE VVQRAYRAIR HIHRRIRQGL ERILLQVHAS
                     *                                                  ├linker

RVIN.
```

FIGURE 7

FIGURE 3



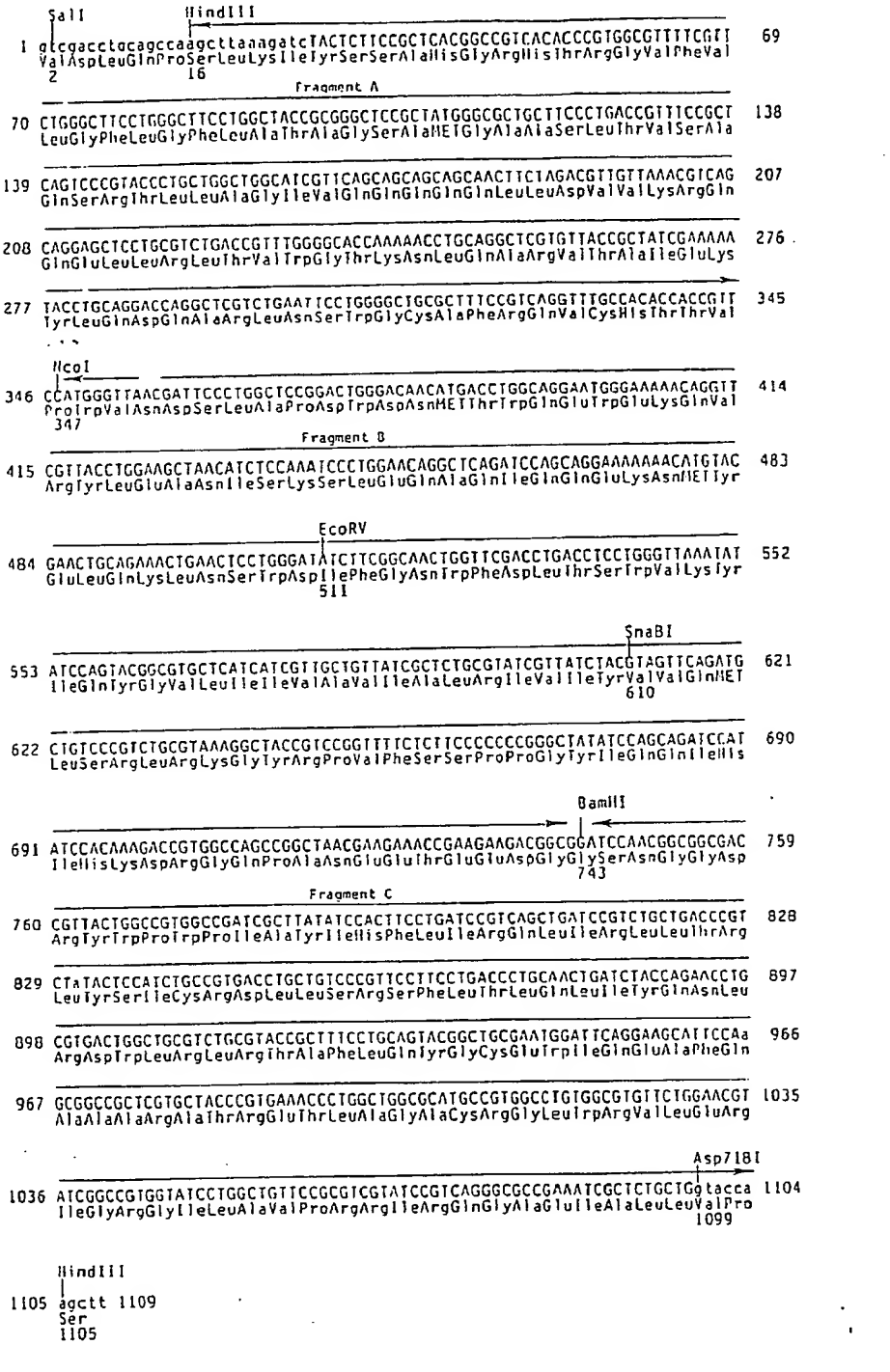


FIGURE 9

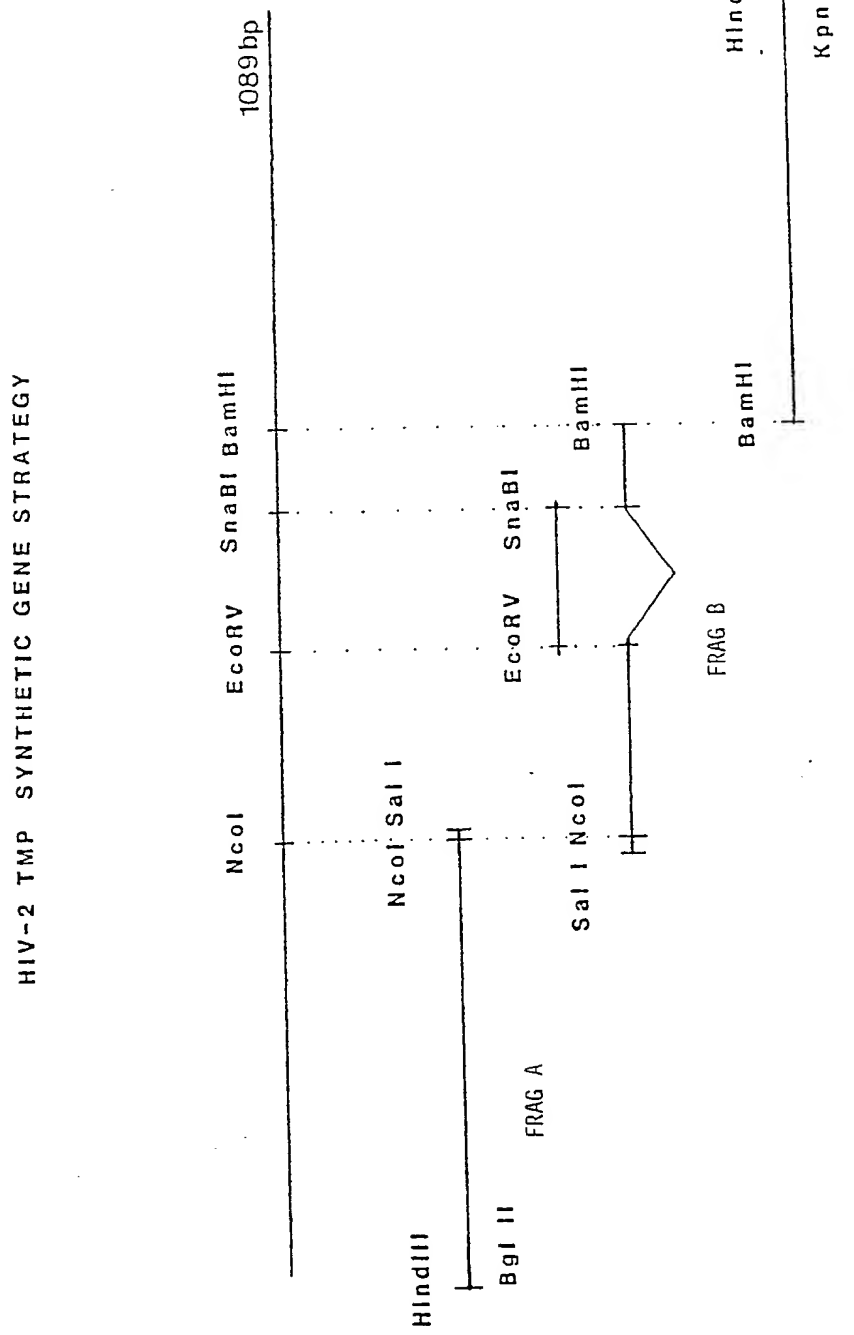


FIGURE 10

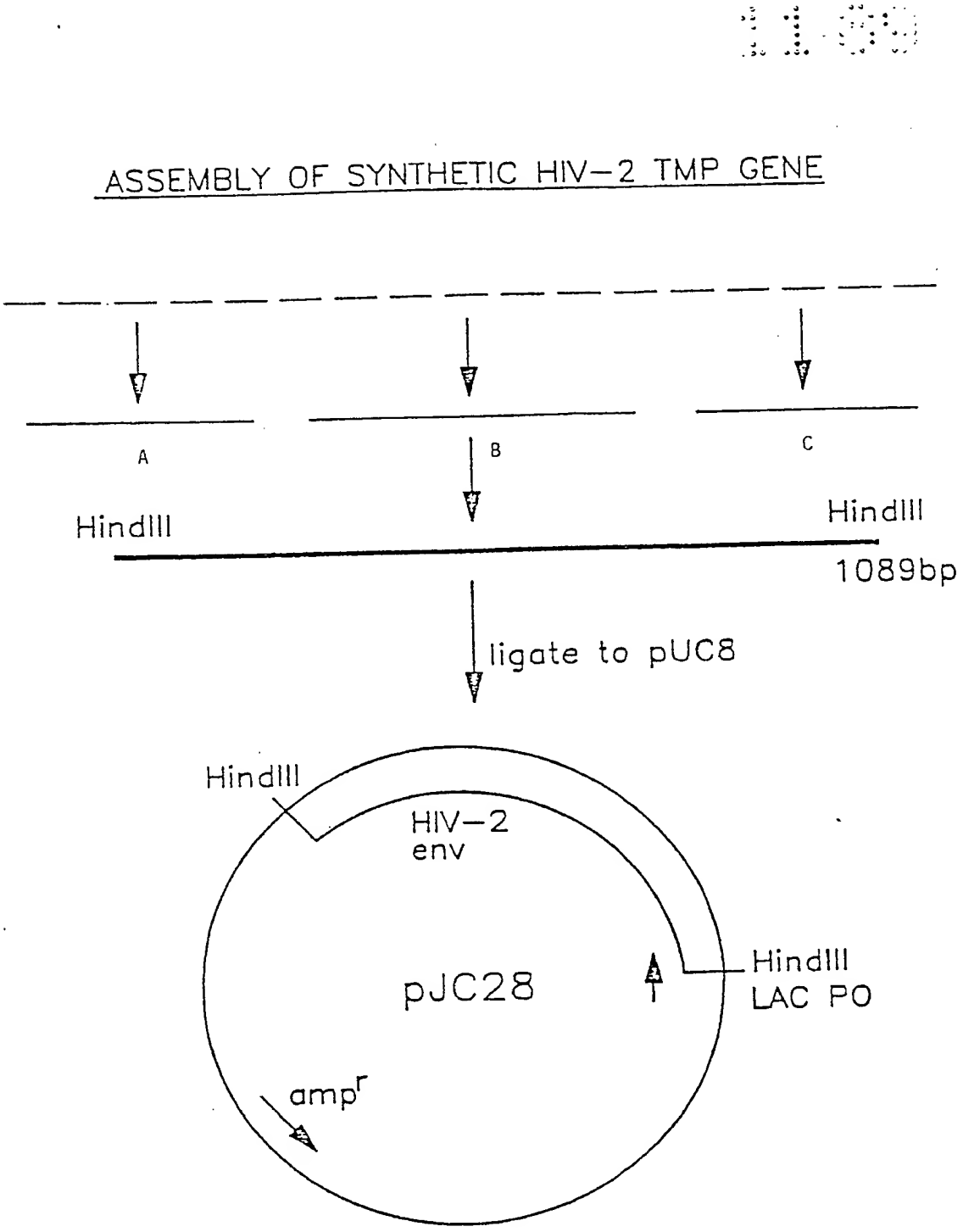

ASSEMBLY OF SYNTHETIC HIV-2 TMP GENE
A
B
C
HindIII
HindIII
1089bp
ligate to pUC8
HindIII
HIV-2
env
pJC28
HindIII
LAC PO
amp^r

FIGURE 11

SYNTHETIC HIV-2 TMP CLONING STRATEGY

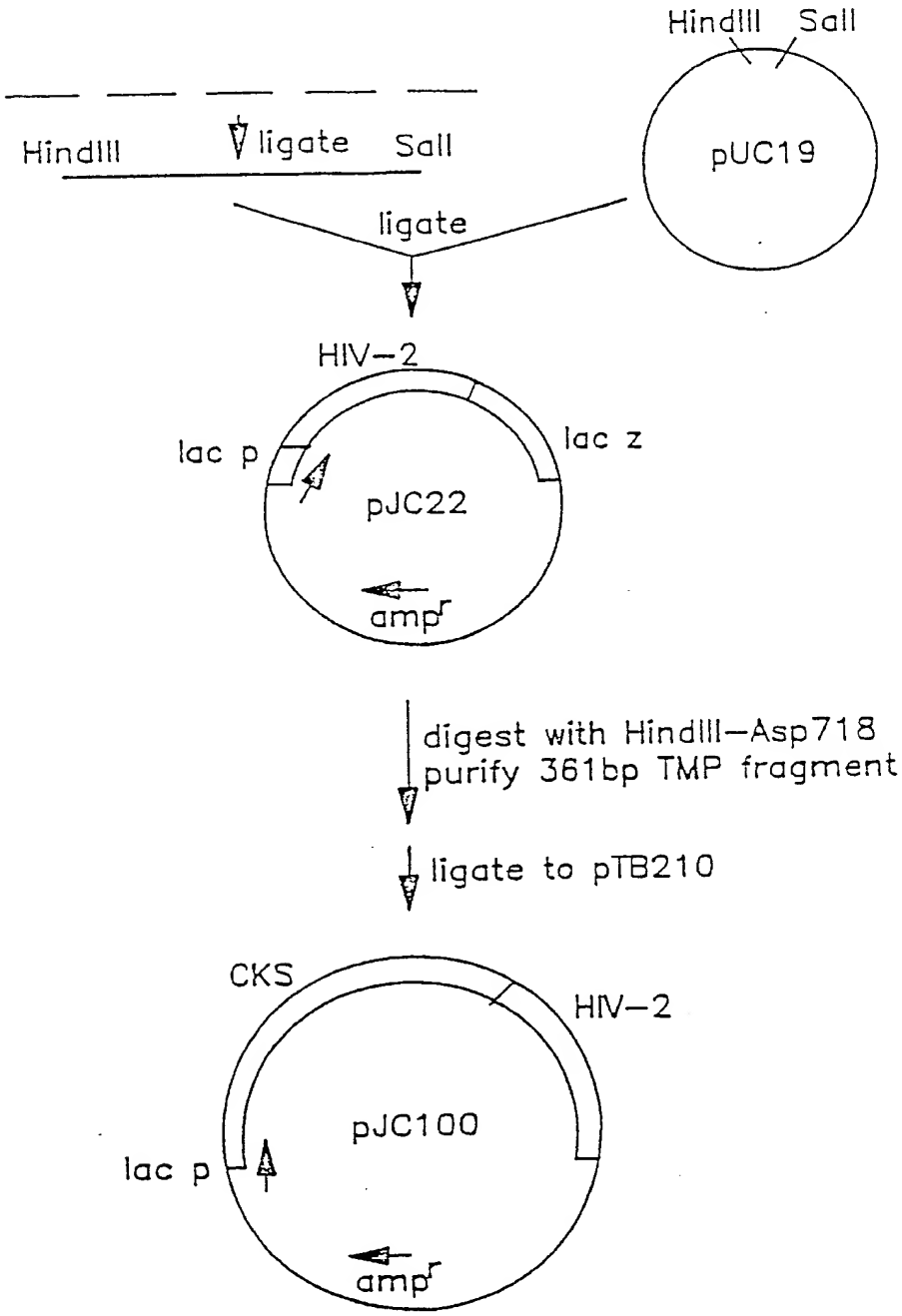


FIGURE 12

EXPRESSION OF SYNTHETIC HIV-2 TMP

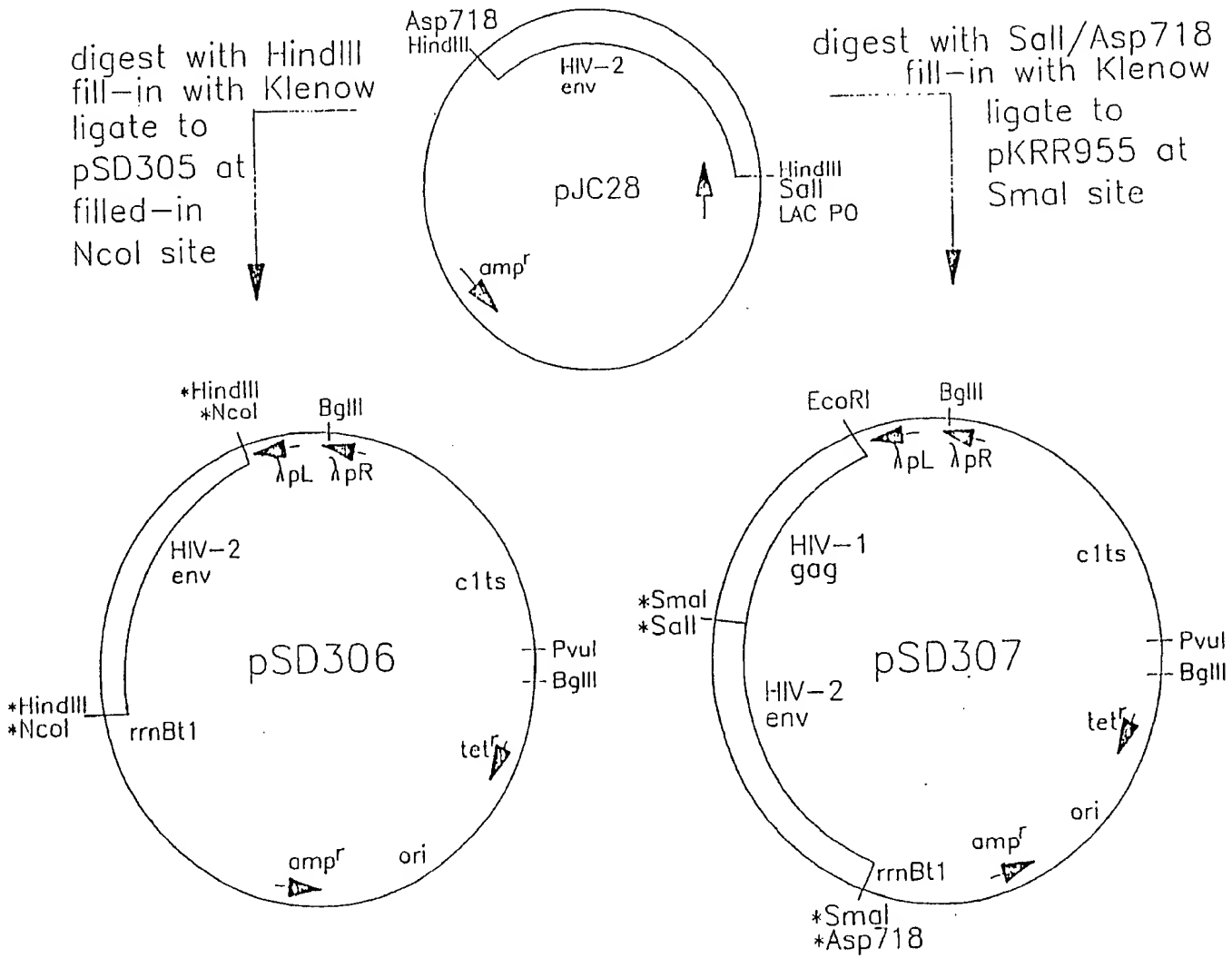


FIGURE 13

EP 0 370 458 A2

PSD306.PEP

```
        10          20          30          40          50          60          70
MSLKIYSSAH GRHTRGVFVL GFLGFLATAG SAMGAASLTV SAQSRTLLAG IVQQQQQLLD VVKRQQELLR
linker ├ HIV-2 TMP seq ──►
        80          90         100         110         120         130         140
LTVWGTKNLQ ARVTAIEKYL QDQARLNSWG CAFRQVCHTT VPWVNDSLAP DWDNMTWQEW EKQVRYLEAN

       150         160         170         180         190         200         210
ISKSLEQAQI QQEKNMYELQ KLNSWDIFGN WFDLTSWVKY IQYGVLIIVA VIALRIVIYV VQMLSRLRKG

       220         230         240         250         260         270         280
YRPVFSSPPG YIQQIHIHKD RGQPANEETE EDGGSNGGDR YWPWPIAYIH FLIRQLIRLL TRLYSICRDL

       290         300         310         320         330         340         350
LSRSFLTLQL IYQNLRDWLR LRTAFLQYGC EWIQEAFQAA ARATRETLAG ACRGLWRVLE RIGRGILAVP

       360         370
RRIRQGAEIA LLVPSSWQFG PG.
          │linker ──
```

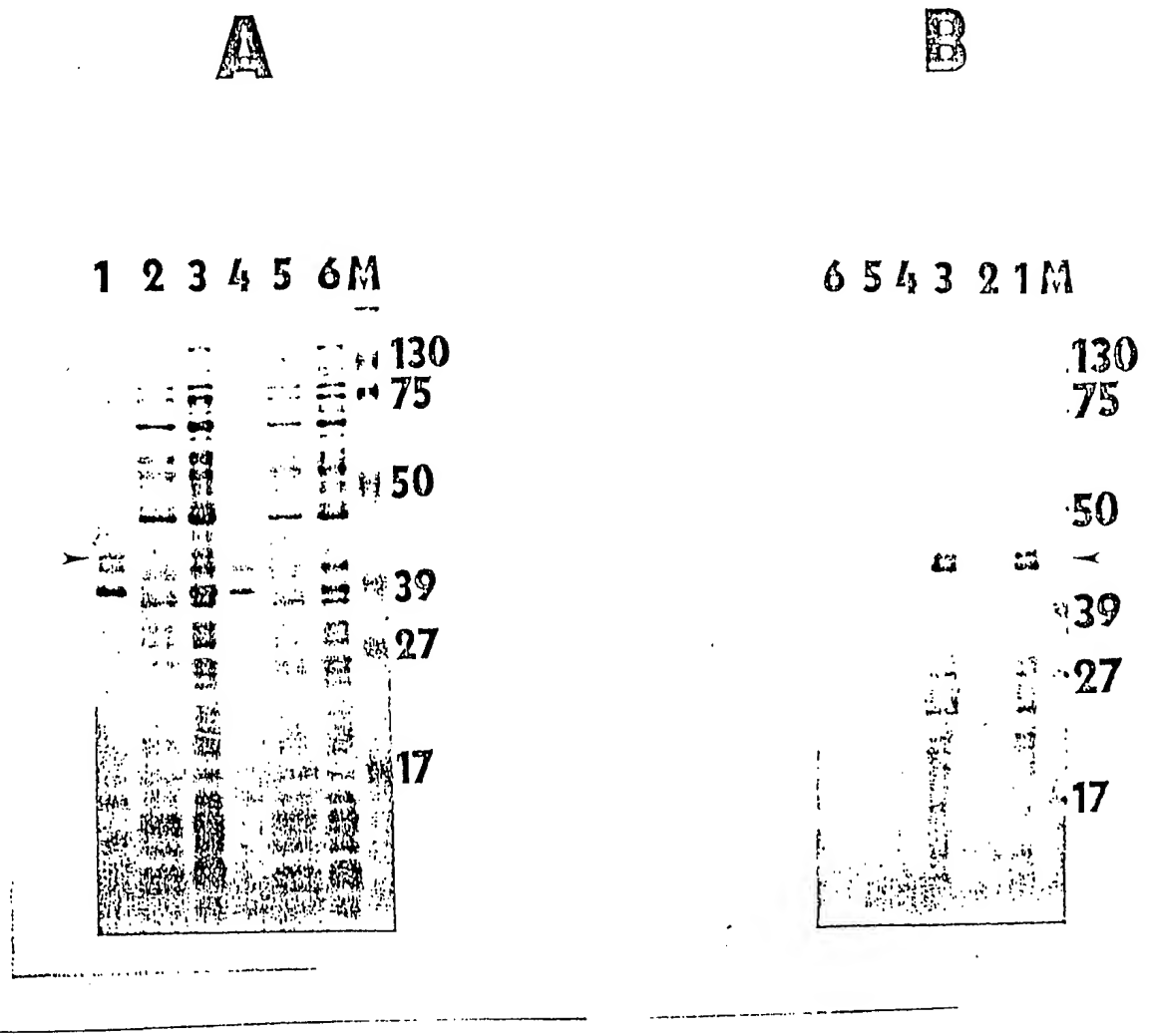
PSD307.PEP

```
        10          20          30          40          50          60          70
MTMITPSLAA GPDTGHSSQV SQNYPIVQNI QGQMVHQAIS PRTLNAWVKV VEEKAFSPEV IPMFSALSEG
linker seq ──────┼──HIV-1 gag seq ──►
        80          90         100         110         120         130         140
ATPQDLNTML NTVGGHQAAM QMLKETINEE AAEWDRVHPV HAGPIAPGQM REPRGSDIAG TTSTLQEQIG

       150         160         170         180         190         200         210
WMTNNPPIPV GEIYKRWIIL GLNKIVRMYS PTSILDIRQG PKEPFRDYVD RFYKTLRAEQ ASQEVKNWMT

       220         230         240         250         260         270         280
ETLLVQNANP DCKTILKALG PAATLEEMMT ACQGVGGPGH KARVLAEAMS QVTNTATIMM QRGNFRNQRK

       290         300         310         320         330         340         350
MVKCFNCGKE GHTARNCRAL DLQPSLKIYS SAHGRHTRGV FVLGFLGFLA TAGSAMGAAS LTVSAQSRTL
                    └─ linker ──┼── HIV-2 TMP seq ──►
       360         370         380         390         400         410         420
LAGIVQQQQQ LLDVVKRQQE LLRLTVWGTK NLQARVTAIE KYLQDQARLN SWGCAFRQVC HTTVPWVNDS

       430         440         450         460         470         480         490
LAPDWDNMTW QEWEKQVRYL EANISKSLEQ AQIQQEKNMY ELQKLNSWDI FGNWFDLTSW VKYIQYGVLI

       500         510         520         530         540         550         560
IVAVIALRIV IYVVQMLSRL RKGYRPVFSS PPGYIQQIHI HKDRGQPANE ETEEDGGSNG GDRYWPWPIA

       570         580         590         600         610         620         630
YIHFLIRQLI RLLTRLYSIC RDLLSRSFLT LQLIYQNLRD WLRLRTAFLQ YGCEWIQEAF QAAARATRET

       640         650         660         670
LAGACRGLWR VLERIGRGIL AVPRRIRQGA EIALLVRVIN .
                                    ├─linker
```

PEP:

FIGURE 14

FIGURE 15



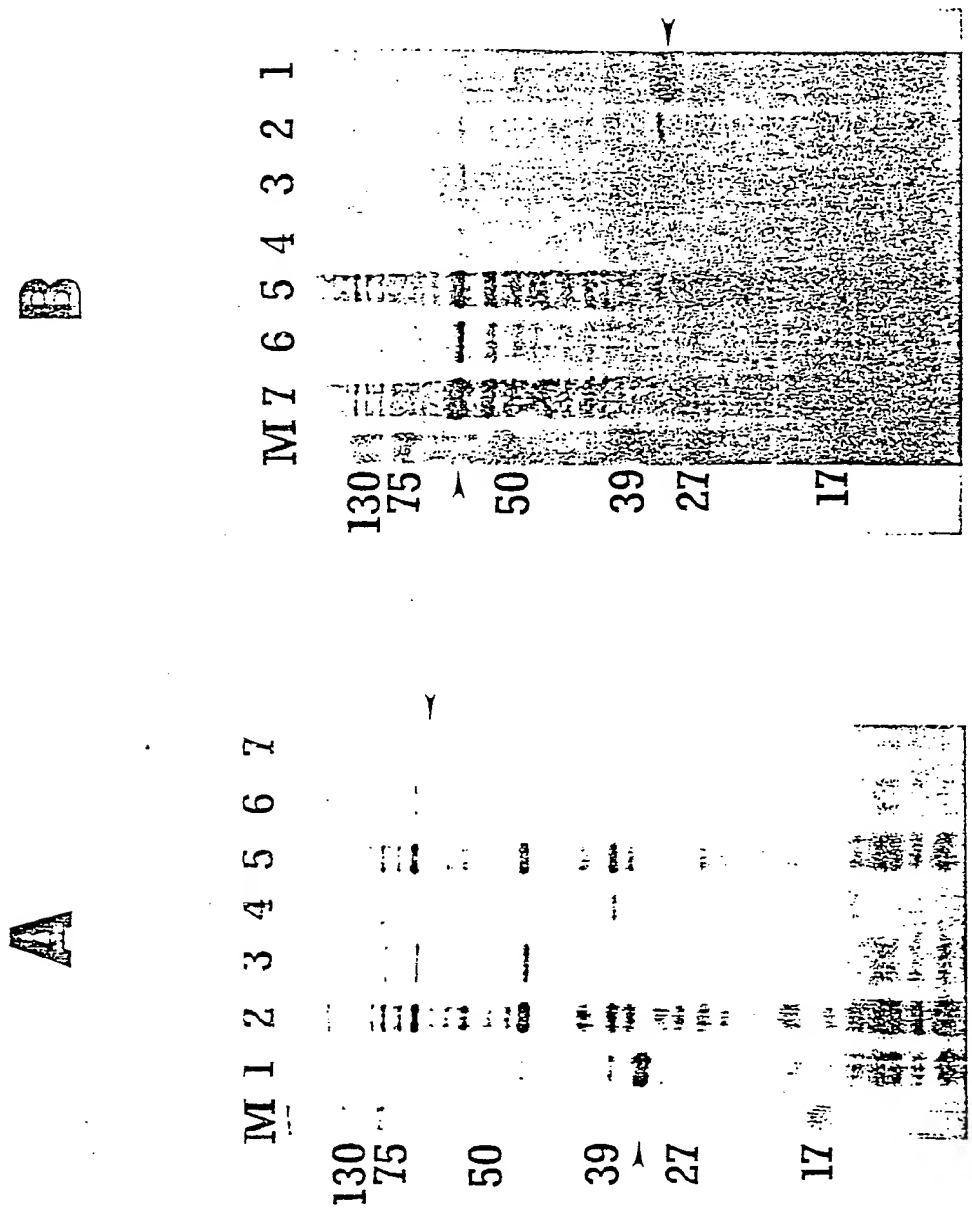


FIGURE 16

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 370 458 A3**

(12) # EUROPEAN PATENT APPLICATION

(21) Application number: **89121513.9**

(22) Date of filing: **21.11.89**

(51) Int. Cl.5: **C07K 13/00**, C12N 15/62, G01N 33/569

(30) Priority: **23.11.88 US 275309**

(43) Date of publication of application:
30.05.90 Bulletin 90/22

(84) Designated Contracting States:
AT BE CH DE ES FR GB GR IT LI NL SE

(88) Date of deferred publication of the search report:
06.11.91 Bulletin 91/45

(71) Applicant: **ABBOTT LABORATORIES**
CHAD-0377, AP6D/2, One Abbott Park Road
Abbott Park, Illinois 60064-3500(US)

(72) Inventor: **Devare, Sushil G.**
2492 Farnsworth Lane
Northbrook Illinois 60062(US)
Inventor: **Suresh, Desai M.**
1408 Amy Lane
Libertyville Illinois 60048(US)
Inventor: **Casey, James M.**
4567 McClure
Gurnee Illinois 60031(US)

(74) Representative: **Modiano, Guido et al**
MODIANO, JOSIF, PISANTY & STAUB
Modiano & Associati Via Meravigli, 16
I-20123 Milano(IT)

(54) **Synthetic DNA derived recombinant HIV antigens.**

(57) The present invention provides a method of synthesizing genes encoding unique HIV-1 and HIV-2 envelope proteins and their fragments, thereby allowing overexpression of these proteins in *E. coli*. The HIV envelope proteins and their fragments have been expressed at high levels as individual proteins or in fusion with other proteins. The HIV envelope proteins thus expressed in *E. coli* can be effectively used for the detection of exposure to HIV as well as the discrimination of HIV-1 and HIV-2.

EP 0 370 458 A3

Clustered order of selected sequences:

2.	CDC42FRAG.PEP	(1-107)
3.	8H1O2FRAG.PEP	(1-107)
4.	SF2FRAG.PEP	(1-107)
1.	MALFRAG.PEP	(1-107)
5.	SYNFRAG.PEP	(1-107)

```
2    1 kAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGfWGCSGKLICTTAVPWNASWSNKtLdQIWNNMT
3    1 EAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWNNMT
4    1 EAQQHLLQLTVWGIKQLQARVLAVERYLrDQQLLGiWGCSGKLICTTAVPWNASWSNKSLEDIWdNMT
1    1 EAQQHLLQLTVWGIKQLQARVLAVERYLqDQrLLGmWGCSGKhICTTfVPWNsSWSNrSLdQIWnNMT
5    1 KAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEDIWNNMT

2   69 WMEWDREIdNYThLIytLIEESQNQQEKNqQELLqLDKW
3   69 WMEWDREInNYTsLIhsLIEESQNQQEKNEQELLELDKW
4   69 WMQWEREIdNYTntIYtLIEESQNQQEKNEQELLELDKW
1   69 WMQWEkEIsNYTgiIYnLiEESQiQQEKNEkELLELDKW
5   69 WMQWEREINNYTNLIYSLLEESQNQQEKNEQELLQLDKW
```

FIGURE 1

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 89 12 1513

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
Y	EP-A-0 187 041 (GENENTECH) * Claims * ---	1,4,7,8	C 07 K 13/00 C 12 N 15/62 G 01 N 33/569
Y	EP-A-0 199 301 (HOFFMANN-LA ROCHE) * Whole document * ---	1,4-8, 13-16	
Y	EP-A-0 001 931 (GENENTECH) * Whole document * ---	7,8	
Y	GENE, vol. 45, no. 3, 1986, pages 317-325, Elsevier Science B.V. (Biomedical Division), Amsterdam, NL; K.A. KELLEY et al.: "Synthesis of fusion and mature murine alpha interferons in Escherichia coli" * Whole article * ---	1,4,7,8	
Y	THE EMBO JOURNAL, vol. 6, no. 3, March 1987, pages 591-598, Eynsham, Oxford, GB; H. WEBER et al.: "Single amino acid changes that render human IFN-alpha 2 biologically active on mouse cells" * Whole article, specially table II * ---	1,4,7,8	TECHNICAL FIELDS SEARCHED (Int. Cl.5) C 12 N C 07 K
P,Y	EP-A-0 331 961 (ABBOTT LABORATORIES) * Whole document * ---	5,6	
Y	WO-A-8 803 562 (OXFORD GENE SYSTEMS LTD) * Whole document * ---	1,4,7,8 ,12-16	
D,Y	NATURE, vol. 326, 16th April 1987, pages 662-669; M. GUYADER et al.: "Genome organization and transactivation of the human immmunodeficiency virus type 2" * The whole document * --- -/-	2-6,9, 10-13, 15,16	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	13-03-1991	CHAMBONNET F.J.

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P0401)

European Patent Office

CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing more than ten claims.

☐ All claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for all claims.

☐ Only part of the claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims and for those claims for which claims fees have been paid.

namely claims:

☐ No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims.

x LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirement of unity of invention and relates to several inventions or groups of inventions.

namely:

1. Claims 1,7,8 and 4-6,13-16(partially): Synthetic HIV1 env antigen; acid nucleic sequence, peptide, fusion-protein and use in diagnostics.

2. Claims 2,3,9-12 and 4-16,13-16(partially): Synthetic HIV2 env antigen: acid nucleic sequences, peptides, fusion-peptides and use in diagnostics.

☒ All further search fees have been paid within the fixed time limit. The present European search report has been drawn up for all claims.

☐ Only part of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid.

namely claims:

☐ None of the further search fees has been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims.

namely claims:

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 89 12 1513

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
P,Y	EP-A-0 400 245 (INSTITUT PASTEUR) * Table 1; claims; figures * -----	2-6,10-12,15,16	

TECHNICAL FIELDS SEARCHED (Int. Cl.5)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	13-03-1991	CHAMBONNET F.J.

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P0401)